PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 19/00, C12P 19/34, C12Q 1/68

A1

(11) International Publication Number: WO 99/06423

(43) International Publication Date: 11 February 1999 (11.02.99)

(21) International Application Number: PCT/US98/15949

(22) International Filing Date: 29 July 1998 (29.07.98)

(30) Priority Data:

60/054,234	30 July 1997 (30.07.97)	US
60/054,209	30 July 1997 (30.07.97)	US
60/054,212	30 July 1997 (30.07.97)	US
60/054,218	30 July 1997 (30.07.97)	US
60/054,214	30 July 1997 (30.07.97)	US
60/054,236	30 July 1997 (30.07.97)	US
60/054,215	30 July 1997 (30.07.97)	US
60/054,211	30 July 1997 (30.07.97)	US
60/054,217	30 July 1997 (30.07.97)	US
60/054,213	30 July 1997 (30.07.97)	US
60/055,968	18 August 1997 (18.08.97)	US
60/055,969	18 August 1997 (18.08.97)	US
60/055,972	18 August 1997 (18.08.97)	US
60/056,561	19 August 1997 (19.08.97)	US
60/056,534	19 August 1997 (19.08.97)	US
60/056,729	19 August 1997 (19.08.97)	US
60/056,543	19 August 1997 (19.08.97)	US
60/056,727	19 August 1997 (19.08.97)	US
60/056,730	19 August 1997 (19.08.97)	US
60/056,554	19 August 1997 (19.08.97)	US

(71) Applicant *(for all designated States except US)*: HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). FENG, Ping [CN/US]; 4 Relda Court, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 1615 Q. Street, N.W. #807, Washington, DC 20009 (US). MOORE, Paul, A. [US/US]; Apartment 104, 1908 Holly Ridge Drive, McLean, VA 22102 (US). SHI, Yanggu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). KYAW, Hla [CN/US]; 520 Sugarbush Circle, Frederick, MD 21703 (US). LI, Yi [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). ZENG, Zhizhen [CN/US]; 13950 Saddleview Drive, Gaithersburg, MD 20878 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). WEI, Ying-Fei [CN/US]; 13524 Straw Bale Lane, Darnestown, MD 20878 (US). FAN, Ping [CN/US]; Apartment 302, 335 West Side Drive, Gaithersburg, MD 20878 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US).

(74) Agents: HOOVER, Kenley, K. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.
With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.

(54) Title: 83 HUMAN SECRETED PROTEINS

(57) Abstract

The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

83 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoeitin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that

encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein , a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide"

refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH_2PO_4; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking

reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched , for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural

processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

This gene is expressed in a broad variety of tissues and cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, liver cancer. Similarly, polypeptides and antibodies directed to

these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., hepatic, cancerous and wounded tissues) or bodily fluids (e.g., bile, lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:99 as residues: Ser-34 to Arg-39, Leu-50 to Ser-55.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1747 of SEQ ID NO:11, b is an integer of 15 to 1761, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

Preferred polypeptides encoded by this gene comprise the following amino acid sequence: MEQTWTRDYFAEDDGEMVPRTSHTAAFLSDTKDRGPPVQSQIWRSGEKVPFVQTYSLRAFEKPPQVQTQALRDFEKHLNDLKKENFSLKLXIYFLEERMQQKYEASREDIYKRNTELKVEVESLKRELQDKKQHLDKTWADVENLNSQNEAELRRQFEERHXETEHVYELLENKXQLLQEESRLAKNEAARMAALVEAEKECNLELSEKLKGVTKNWEDVPGDQVKPDQYTEALAQRDK (SEQ ID NO:188) or MVPRTSHTAAFLSDTKDRGPPVQSQIWRSGEKVPFVQTYSLRAFEKPPQVQTQALRDFEKHLNDLKKENFSLKLXIYFLEERMQQKYEASREDIYKRNTELKVEVESLKRELQDKKQHLDKTWADVENLNSQNEAELRRQFEERHXETEHVYELLENKXQLLQEESRLAKNEAARMAALVEAEKECNLELSEKLKGVTKNWE

DVPGDQVKPDQYTEALAQRDK (SEQ ID NO:187). Polynucleotides encoding these polypeptides are also provided.

This gene is expressed primarily in infant brain and to a lesser extent in a large variety of other tissues, organs and cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, various forms of congentital mental retartation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., developmental, cancerous and wounded tissues) or bodily fluids (e.g., amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:100 as residues: Met-1 to Arg-22, Leu-46 to Arg-52, Asn-64 to Gln-70.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1505 of SEQ ID NO:12, b is an integer of 15 to 1519, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

Preferred polypeptide encoded by this gene comprise the following amino acid sequence:
IRHELLPALHLQAHDAAYNLLFFASGGGKFNYQGTKRWLEDNLDHTGERP

RVGVGVPRWWCRGEAXRPRGCHGGSQEAQREGRGPLPGPHPPRQLSVSC RLQPASGQCGLRAVPGHRGPGQQPAPAXVRPXREGTLQHAFXRELETVAA HQFPEVRFSMVHKRINLAEDVLAWEHERFAIRRLPAFTLSHLESHRDGQRS SIMDVRSRVDSKTLIRLPQPPKVLGLRV (SEQ ID NO:189). Polynucleotides encoding such polypeptides are also provided. This gene is believed to reside on chromosome 19. Thus, polynucleotides related to this gene are useful as chromosome markers in linkage analysis for chromosome 19.

This gene is expressed primarily in microvascular endothelial cells and to a lesser extent in in a variety of other cell types including activated T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, blood circulatory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the circulatory system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., vascular, and blood, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1057 of SEQ ID NO:13, b is an integer of 15 to 1071, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

The gene encoding the disclosed cDNA sequence is believed to reside on chromosome 22. Accordingly, polynucleotides related to this invention are useful in linkage analysis as markers for chromosome 22.

This gene is expressed primarily in a variety of cell types of muscle and bone origin.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, osteoporosis or any of a variety of disease that involve wasting to bone or muscle. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal and muscular systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., musculoskeletal, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:102 as residues: Lys-81 to Thr-92, Arg-168 to Tyr-176, Gly-199 to Ser-216.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 941 of SEQ ID NO:14, b is an integer of 15 to 955, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

This gene is expressed primarily in the brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, various brain disorders including mood disorders, memory disorders, depression, and seizures. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., neural, cancerous and wounded tissues) or bodily fluids (e.g., lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:103 as residues: Ser-62 to Cys-67.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1494 of SEQ ID NO:15, b is an integer of 15 to 1508, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

When tested against PC12 cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response gene 1) pathway. Thus, it is likely that this gene activates sensory neuron cells through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jaks-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation. This gene maps to chromosome 1, and therefore, may be used as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in small intestine, and to a lesser extent, in fetal liver and infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, intestinal cancers, premalignancies, and development disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.gastrointesinal, developing, or cancerous and wounded tissues) or bodily fluids (e.g.amniotic fluid, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of intestinal cancers and premalignancies, or ulcers, intestinal infections or other conditions arising from disorders of the gastrointesinal system. Alternatively, based upon the detected EGR activity in sensory neurons may suggest that the protein product of this gene is useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one

or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1992 of SEQ ID NO:16, b is an integer of 15 to 2006, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

The translation product of this gene was shown to have homology to the human LAK-4p which is thought to be involved in T-cell activation as this gene is specifically expressed during such a response (See Genebank Accession No.gnl|PID|d1025089 (AB002405)). Preferred polypeptides comprise the following amino acid sequence: IYLNIQVVRGQRKVICLLKEQISNEGEDKIFLINKLHSIY (SEQ ID NO:190), ERKEREERSRVGTTEEAAAPPALLTDE (SEQ I D NO:191), and/or RHEMENT (SEQ ID NO:192),. Also preferred are the polynucleotides encoding these polypeptides.

This gene is expressed primarily in several types of leukocytes, thymus, bone marrow, and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune system disorders, particularly of the leukocytes. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.immune, hematopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:105 as residues: Gln-38 to Asp-45, Glu-58 to Arg-67.

The protein product of the gene, based upon its homology to the human immune-specific LAK-4p protein, in addition to its tissue distribution in leukocytes, is likely to be a modulator of the immune system and could be used in a variety of

theraputic situations which require modulation of immune cell production such as leukemias and in protection of hematoprogenitors during chemotherapy. Additionally, the protein product of this gene is useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in thymus and bone marrow indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 531 of SEQ ID NO:17, b is an integer of 15 to 545, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

This gene is expressed primarily in lymphocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but

are not limited to, diseases of the immune system, particularly of the lymphocytes. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.immune, hematopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in lymphocytes indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 588 of

SEQ ID NO:18, b is an integer of 15 to 602, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

This gene is expressed primarily in the human embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, dvelopmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.developing, differentiating, or cancerous and wounded tissues) or bodily fluids (e.g.amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 573 of

SEQ ID NO:19, b is an integer of 15 to 587, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

This gene is expressed primarily in the human embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, reproductive and developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.developing, differentiating, or cancerous and wounded tissues) or bodily fluids (e.g.amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:108 as residues: Asn-6 to Ser-13.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present

invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 630 of SEQ ID NO:20, b is an integer of 15 to 644, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

This gene is expressed primarily in the human embryo and the prostate.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, developmental and reproductive disorders, particularly with prostate cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive and urogenital systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.developmental, reproductive, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Alternatively, expression within the prostate indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of prostate cancer, and related reproductive disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1243 of SEQ ID NO:21, b is an integer of 15 to 1257, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

This gene is expressed primarily in the human embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, Developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.developmental, or cancerous and wounded tissues) or bodily fluids (e.g. amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:110 as residues: Trp-6 to Arg-13.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the

above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 527 of SEQ ID NO:22, b is an integer of 15 to 541, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

This gene is expressed primarily in the human embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, reproductive and developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.developmental, or cancerous and wounded tissues) or bodily fluids (e.g.amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the

protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 553 of SEQ ID NO:23, b is an integer of 15 to 567, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

This gene maps to chromosome 11, and therefore, may be used as a marker in linkage analysis for chromosome 11.

This gene is expressed primarily in immune cells, particularly T cells and dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune disorders, particularly immunodeficiences such as AIDS. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.immune, hematopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation

of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 572 of SEQ ID NO:24, b is an integer of 15 to 586, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:24, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

This gene maps to chromosome 1, and therefore, may be used as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in the brain and, to a lesser extent, in prostate and kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, disorders of the brain and CNS, such as Alzheimer's and Parkinson's disease. Similarly, polypeptides and antibodies directed to these

polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, urogenital, or cancerous and wounded tissues) or bodily fluids (e.g.seminal fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, the tissue distribution in kidney indicates that this gene or gene product could be used in the treatment and/or detection of kidney diseases including renal failure, nephritus, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilms Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1496 of SEQ ID NO:25, b is an integer of 15 to 1510, where both a and b correspond to the positions of nucleotide

residues shown in SEQ ID NO:25, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

This gene was found to have homology to both the human ni06c07.s1 and mouse Mpgc60 cDNAs which are specifically expressed in intestinal tissue (See Genebank Accession Nos AA526969 and gb|Y11505|MMMPGC60, respectively). As such, it is probable that the translation product of this gene is useful for the diagnosis, treatment, and/or prevention of various gastrointestinal disorders and afflictions.

This gene is expressed primarily in multiple tissues, including the brain, breast, and kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, disorders involving the brain and central nervous system, such as Alzheimer's and Parkinson's, and reproductive and gastrointestinal disorders. Also disorders of the breast and kidney, including cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and central nervous system also the urogenital system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, endothelial, hepatic, and mammary, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, bile, breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:114 as residues: Pro-3 to Pro-9.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors,

including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, considering the homology to intestinal-specific proteins may suggest that the translation product of this gene is useful for the diagnosis, treatment, and/or prevention of various gastrointestinal disorders.Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 521 of SEQ ID NO:26, b is an integer of 15 to 535, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

This gene maps to chromosome 19, and therefore, may be used as a marker in linkage analysis for chromosome 19.

This gene is expressed in multiple tissues, particularly brain and placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, disorders of the brain and central nervous system, such as Alzheimer's and Parkinson's, in addition to reproductive and developmental disorders. Also disorders of the reproductive system Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, central nervous system, and the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, reproductive, or cancerous and wounded tissues) or bodily fluids (e.g.amniotic

fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:115 as residues: Pro-6 to Glu-35, Ser-47 to Glu-52, Gly-67 to Trp-73, Arg-85 to Asn-90, Asn-114 to Asn-119, Thr-134 to Ser-141, Asn-250 to Glu-260.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, the tissue distribution in placenta indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders since development relies on decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1259 of SEQ ID NO:27, b is an integer of 15 to 1273, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

The translation product of this gene was found to have homology to several collagen proteins.

This gene is expressed primarily in cells of the immune system, including monocytes and neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, disorders affecting the immune systems such as AIDS and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.immune, hematopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.EGE, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in immune cells combined with its homology to collagen would suggest that this protein may also be important in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid, and would healing disorders. Alternatively, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils and monocytes indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including

arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 766 of SEQ ID NO:28, b is an integer of 15 to 780, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

This gene is expressed primarily in hepatocellular tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hepatoma, and other disorders of the liver. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.hepatic, or cancerous and wounded tissues) or bodily fluids (e.g.bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those

comprising a sequence shown in SEQ ID NO:117 as residues: Glu-33 to Glu-56, Thr-75 to Cys-81.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). The tissue distribution in hepatic tumors indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of these tumors, in addition to other tumors where expression has been indicated. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 805 of SEQ ID NO:29, b is an integer of 15 to 819, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

This gene is expressed primarily in apoptotic T cell.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hematopoietic and immune diseases, or cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.hematopoietic, immune or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having

such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 594 of SEQ ID NO:30, b is an integer of 15 to 608, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

The translation product of this gene shares sequence homology with mouse erythroid ankyrin protein which is thought to be important in linking the spectrin-based membrane skeleton to the plasma membrane in red blood cells. As such, the

translation product of this gene may show utility in the treatment and/or diagnosis of various hematopoietic disorders involving structural anomalies such as thalassemia and sickle-cell anemia syndromes (See Genebank Accession No. gil311822). When tested against K562 cell lines, supernatants removed from cells containing this gene activated the ISRE (interferon-sensitive responsive element) pathway. Thus, it is likely that this gene activates kindey cells through the Jaks-STAT signal transduction pathway. The ISRE is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in colon cancer cells and, to a lesser extent, in pancreatic and testical tumors.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, cancers and tumors of the urogenital, hematopoietic, or endocrine systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the digestive and repruductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.hematopoietic, urogenital, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:119 as residues: Met-1 to Gly-6, Lys-13 to Tyr-18, Asp-23 to Asp-28, Leu-55 to Glu-60, Pro-148 to Gly-155.

The tissue distribution combined with the observed ISRE activity indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within tumor tissues and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Additionally, the homology to a structural protein in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or

survival of hematopoietic cell lineages. In such an event, this gene may be useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Alternatively, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various endocrine disorders and cancers, particularly Addisonís disease, Cushingís Syndrome, and disorders and/or cancers of the pancrease (e.g. diabetes mellitus), adrenal cortex, ovaries, pituitary (e.g., hyper-, hypopituitarism), thyroid (e.g. hyper-, hypothyroidism), parathyroid (e.g. hyper-,hypoparathyroidism) , hypothallamus, and testes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1203 of SEQ ID NO:31, b is an integer of 15 to 1217, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

This gene maps to chromosome 19, and therefore, may be used as a marker in linkage analysis for chromosome 19.

This gene is expressed primarily in umbilical vein endothelial cells and, to a lesser extent, in human adipose.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions:reproductive or metabolic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential

identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.reproductive, or cancerous and wounded tissues) or bodily fluids (e.g.amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within embryonic tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus, this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Alternatively, expression in adipose tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, prevention, and/or treatment of various metabolic disorders such as Tay-Sachs disease, phenylkenonuria, galactosemia, porphyrias, Hurler's syndrome, or disorders related to lipid metabolism. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 751 of SEQ ID NO:32, b is an integer of 15 to 765, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

This gene is expressed primarily in bone marrow stromal cells, and, to a lesser extent, in epithelial-TNF alpha induced cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, integumentary and hematopoietic diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.immune, hematopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoetic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, expression in cells induced by epithelial TNF-alpha indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within differentiating tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Additionally, the expression in hematopoietic cells and tissues indicates that this protein may play a role in the proliferation, differentiation, and/or survival of hematopoietic cell lineages. In such an event, this gene may be useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could

again be useful in cancer therapy. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 738 of SEQ ID NO:33, b is an integer of 15 to 752, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

This gene maps to chromosome 1, and therefore, may be used as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in brain, and, to a lesser extent, in ovary and activated T-cell.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions:immune deficiencies and brain degenerative diseases, in addition to reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and neural system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, and reproductive, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:122 as residues: Glu-2 to Glu-13, Pro-23 to Cys-36, Glu-47 to Ser-56, Val-64 to Pro-69, Val-106 to Asn-113, Ser-128 to Ala-134, Ser-155 to Thr-163, Lys-176 to Phe-188, Leu-192 to Asp-207, Leu-209 to Gly-232, Glu-262 to Asn-269, Thr-274 to Lys-279, Lys-284 to Gly-294, Pro-309 to Cys-314, Phe-318 to Lys-337.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, female reproductive disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2251 of SEQ ID NO:34, b is an integer of 15 to 2265, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 25

The translation product of this gene shares sequence homology with a mouse fat-specific protein FSP27 which is thought to be important in adipose differentiation (See Genebank Accession No. pir|A42445|A42445). One embodiment of this gene comprises polypeptides of the following amino acid sequence: RKLSTGPFSACKPRATCCFTSCYLQQLLDATEDGHPPKGKASSLIPTCLKILQ (SEQ ID NO:193), TSCYLQQLLDATEDGHPPKGKASSLIPTC (SEQ ID NO:194), and/or CCGAKRIMKEALHWALFSMQATGHV (SEQ ID NO:195). An additional embodiment is the polynucleotides encoding these polypeptides. This gene maps to chromosome 3, and therefore, may be used as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in adipose and to a lesser extent in small intestine and a few other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, adipose related disorders, including lipid metabolism disorders, and obesity. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of adipose tissue, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.adipose and gastrointestinal, or cancerous, or wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:123 as residues: Arg-30 to Gln-41.

The tissue distribution in adipose tissue combined with the homology to ASP27 indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of adipose related disorders, particularly hyper- and hypolidemias, Tay-Sachs, atherosclerosis, and obesity. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 629 of SEQ ID NO:35, b is an integer of 15 to 643, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 26

The translation product of this gene was shown to have homology to a the human KIAA0427 protein, novel, brain-specific protein that may be important in brain development (See Genebank Accession No.gnl|PID|d1025779 (AB007887)). One embodiment of this gene comprises polypeptides of the following amino acid sequence: PPAGATSPGRIIXPXSAVLIPSPVKSYRGWLVMGEPSREEYKIQSFDAETQQ LLKTALKDPGAVDLEKVA NVIVDHSLQDCVFSKEAGRMXYAIIQAESKQAGQSV FRRGLLNRLQQEYQAREQLXARSLQGWVCYVTFICNIFDYLRVNNMPMM ALVNPVYDCLFRLAQPDSLSKEEEVDCLVLQLHRVGEQLEK (SEQ ID NO:196), PGRIIXPXSAVLIPSPVKSYRGWL (SEQ ID NO:197) KQAGQSVFRRGLL NRLQQEYQAREQ (SEQ ID NO:198), and/or YDCLFRLAQPDSLSKEEEVDC (SEQ ID NO:199),. An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in hematopoiesis related tissues and cell types and to a lesser extent in brain and a few cancer cell lines and tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, neural, andinflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, immune, cancerous, or wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:124 as residues: Met-1 to Met-6, Lys-50 to Arg-59.

The tissue distribution in brain combined with the homology to a novel brain-specific human protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic

disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors. including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, the tissue distribution in hematopoetic tissue indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:36 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1288 of SEQ ID NO:36, b is an integer of 15 to 1302, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

The translation product of this gene was shown to have homology to the rat mitochondrial brown-fat uncoupling protein which is an uncoupling protein specific to mitochondrial brown fat and is thought to play an integral role in the thermogenesis of this tissue (See Genebank Accession No.P04633). One embodiment of this gene comprises polypeptides of the following amino acid sequence: MKRTSVNPQTLCEARPAGXSQQPLSLDSEAPRGGVAPPRLQGPPPHQRVHL TLECTTHPTVGKASV LGPCLLLLSCPRAPAGPPPPPHSRVRAGGCRPWARREGH CRPLGADTDTSRICHGRRPFSL (SEQ ID NO:200), MSLPAAPAGRLSPLYWRSS NTRSQLSLLWELGHFFTRCCRRPHPNPHLPALSVCRCHILHKIMLWEPS SPLLPALP (SEQ ID NO:201), and/or MTSPGQGRAGRRGDEGSHNMILCKIWQR HTLRAGRWGLGWGRRQHRVKKCPSSHSKESCDRVFELLQYKGES RPAGAA GRDIIWFP (SEQ ID NO:202). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in hematopoietic tissues and neuronal tissues and to a lesser extent in some cancer and other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, neural, and/or lipid metabolism disorders and/or diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and neuronal tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, adipose, hematopoietic, and cancerous, or wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:125 as residues: Glu-11 to Ser-17.

The tissue distribution in neural tissue combined with the homology to a protein specific to adipose tissue indicates that polynucleotides and polypeptides

corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, preception, and particularly neural disorders involving anomylous lipid metabolism. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, considering the tissue distribution in hematopoetic tissue, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoetic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 991 of SEQ ID NO:37, b is an integer of 15 to 1005, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 28

The translation product of this gene was shown to have homology to the serine protease PfSP6 N-terminal fragment (See Genebank Accession No. W01189) which may show utility in treatment and/or prevention of various insect or worm infestations, and/or diseases. One embodiment of this gene comprises polypeptides of the following amino acid sequence: PSLRGPKAGAPPRWRPL (SEQ ID NO:203), NLVDPPXCRNSARETLKLGRVEVSI (SEQ ID NO:204), KAGAPPR (SEQ ID NO:205), and/or CRNSAR (SEQ ID NO:206). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in breast lymph node and to a lesser extent in some other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, and reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast lymph node, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.immune, reproductive, cancerous, or wounded tissues) or bodily fluids (e.g.lymph, breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in breast lymph nodes indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or

immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 594 of SEQ ID NO:38, b is an integer of 15 to 608, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

This gene is expressed primarily in breast lymph node, and to a lesser extent, in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, and reproductive diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the breast lymph node, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.immune, reproductive, cancerous, or wounded tissues) or bodily fluids (e.g.lymph, breast milk, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:127 as residues: Pro-32 to Gly-39.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in breast lymph nodes indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 911 of SEQ ID NO:39, b is an integer of 15 to 925, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

The translation product of this gene was shown to have homology to the unc-50 related protein of Rattus norvegicus (See Genebank Accession No. gi|2735550) which is thought to be a novel RNA-binding protein that regulates neuronal nicotinic receptor expression.Preferred polypeptides comprise the following amino acid sequence: QDSRKMLPSTSVNSLVQGNG

VLNSRDAARHTAGAKRYKYLRRLFRFRQMDF EFAAWQMLYLFTSPQRVYRNFHYRKQTKDQWARDDPAFLVLLSIWLCV STLGFGFVLD (SEQ ID NO:207) NXQSRDYDVEWGYAFDVHLNAFYPLLV ILHFIQLFFINHVILTDTFIGYLVGNTLWLVAVGYYIYVTFLGYSALPFLKNT VILLYPFAPLILLYGLSLALGWNFTHTLCSFYKYRVK (SEQ ID NO:208), SVNSLVQGNGVLNSRDAARHTAGAKRYKYLRRLFRFRQMDFEFAA (SEQ ID NO:209), VILTDTFIGYLVGNTLWLVAVGY (SEQ ID NO:210), and/or GWNFTHTLCSFYKYRV (SEQ ID NO:211). Also preferred are the polynucleotides encoding these polypeptides.

This gene is expressed primarily in hematopoietic tissues and to a lesser extent in prostate, placenta, and other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune, neural, and inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and neural systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, immune, cancerous, or wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution combined with the homology to a putative, brain-specific transcription factor indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Alternatively, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system

disorders. Expression of this gene product in hematopoeitic tissue indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1205 of SEQ ID NO:40, b is an integer of 15 to 1219, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

This gene is expressed primarily in proliferating tissues and tumors, and to a lesser extent, in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, growth related diseases and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological

probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of tumors such as breast cancer, colon cancer, and many other common cancers expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.differentiating, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders. Expression within tumor tissues and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. In such an event, this gene may be useful in the treatment of lymphoproliferative disorders, and in the maintenance and differentiation of various hematopoietic lineages from early hematopoietic stem and committed progenitor cells. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1710 of SEQ ID NO:41, b is an integer of 15 to 1724, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

This gene is expressed primarily in placenta, lung, and to a lesser extent in other tissues.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, developmental, reproductive, and pulmonary disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the lung and placenta, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.reproductive, pulmonary, or cancerous and wounded tissues) or bodily fluids (e.g.pulmonary surfactant, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:130 as residues: Met-1 to Trp-7, Ala-37 to Arg-48.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of developmental and pulmonary disorders, particularly of cancer since development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 784 of SEQ ID NO:42, b is an integer of 15 to 798, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 33

The translation product of this gene was shown to have homology to the human mitosis-associated nuclear antigen RMSA-1 which may be useful as an antisense therapy for blocking the onset of mitosis (See Genebank Accession No.Q72501).

This gene is expressed primarily in spleen of chronic lymphocytic leukemia.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, disorders of immune system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoetic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and hematopoetic, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in leukemia cells combined with its homology to a mitotic regulatory factor indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or

proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 679 of SEQ ID NO:43, b is an integer of 15 to 693, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

The sequence of this gene was shown to have homology to the guinea pig platelet activating factor (PAF) receptor which is a unique phospholipid mediator, possesses potent proinflammatory, smooth-muscle contractile and hypotensive activities, and appears to be crucial in the pathogenesis of bronchial asthma and in the lethality of endotoxin and anaphylactic shock. Sequence analysis indicates that the receptor belongs to the superfamily of G protein-coupled receptors. (See Genebank Accession No.gb|X56736|CCPAFREC).

This gene is expressed primarily in human brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, disorders related to central nervous system, as well as the hematopoetic system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, hematopeotic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid

from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:132 as residues: Pro-25 to Thr-31.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Considering the homoology to a platelet activating factor, in addition the tissue distribution in brain, indicates that the protein product of this gene may show utility in the diagnosis, treatment, and/or prevention of stroke, amnesia, and other neural disorders related to vascular trauma and inflammation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1344 of SEQ ID NO:44, b is an integer of 15 to 1358, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:44, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 35

This gene is expressed primarily in primary dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, skin disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential

identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the epithelial cells of skin, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.integumentary, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:133 as residues: His-106 to Ser-117.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowenís disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Pagetís disease, mycosis fungoides, and Kaposiís sarcoma), injuries and inflammation of the skin (i.e.wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. Moreover, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, althletes foot, and ringworm). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 951 of SEQ ID NO:45, b is an integer of 15 to 965, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

This gene is expressed primarily in macrophages.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hematopoetic, and/or immune disorders and afflictions, particularly bacterial infections. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immunesystem, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and hematopoetic, immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:134 as residues: Ser-12 to Trp-19, Val-51 to Thr-57, Ser-103 to Glu-116, His-123 to Leu-130, Gln-138 to Gly-143.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against

the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 777 of SEQ ID NO:46, b is an integer of 15 to 791, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 37

This gene is expressed primarily in human microvascular endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, vascular diseases, particularly stroke. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the blood vescle system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.vascular, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis, treatment, and/or prevention of vascular diseases, such as vasculitis, varicose veins, stroke, aneurysm, in addition to disorders involving vasodilation and constriction. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have

been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 756 of SEQ ID NO:47, b is an integer of 15 to 770, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 38

This gene is expressed primarily in human rhabdomyosarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neuromuscular disorders, particularly cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the muscular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.muscle, neural, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:136 as residues: Ser-82 to Val-87, Pro-103 to Gly-110.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various muscle disorders, such as muscular dystrophy, cardiomyopathy, fibroids, myomas, and rhabdomyosarcomas. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present

invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 861 of SEQ ID NO:48, b is an integer of 15 to 875, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

This gene is expressed primarily in spleen metastatic melanoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hematopoetic, and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoeitic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.hematopoetic, immune, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:137 as residues: Met-1 to Lys-7.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well

as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 600 of SEQ ID NO:49, b is an integer of 15 to 614, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

The translation product of this gene was shown to have homology to a zinc finger protein of Rattus norvegicus which is known to be testis-specific and, as such, may suggest that the protein would have utility as a transcription factor (See Genebank Accession No. gi|57504). One embodiment of this gene comprises polypeptides of the following amino acid sequence: PMVLKLKDWPPGEDFRDMMP (SEQ ID NO:212), YFVRPDLGPKMYNAYG (SEQ ID NO:213), NSAREDGQP (SEQ ID NO:214), and/or LNLASRLP (SEQ ID NO:215). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in bone marrow cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and hematopoetic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system,

expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.hematopoetic, immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoetic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, based upon the homology to a testis-specific zinc finger protein may suggest that the protein product of this gene is useful in the diagnosis, treatment, and/or prevention of various male reproductive disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 542 of SEQ ID NO:50, b is an integer of 15 to 556, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

This gene is expressed primarily in bone marrow cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, obesity and bone marrow disorders, particularly of the hematopoetic system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.hematopoetic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:139 as residues: Arg-52 to Asn-60, Asn-65 to Ala-73, Ala-81 to Ser-89.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoetic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 989 of SEQ ID NO:51, b is an integer of

15 to 1003, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

This gene is expressed primarily in teratocarcinoma cells, and to a lesser extent in human amygdala.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neural disorders, particularly cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain or CNS, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:140 as residues: Pro-20 to Cys-26.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, preception, and particularly cancer. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of

the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 872 of SEQ ID NO:52, b is an integer of 15 to 886, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:52, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 43

When tested against promyelocyctic cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activation site) pathway. Thus, it is likely that this gene activates myeloid cells through the Jaks-STAT signal transduction pathway GAS is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in human neutrophil.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hematopoetic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoetic system. expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:141 as residues: Gly-11 to Ser-18, Thr-26 to Lys-36.

The tissue distribution combined with the biological activity within myeloid cells indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders.

Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 550 of SEQ ID NO:53, b is an integer of 15 to 564, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 44

This gene is expressed primarily in human neutrophil.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hematopoetic and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of

disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:142 as residues: Leu-41 to Glu-48.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 919 of SEQ ID NO:54, b is an integer of 15 to 933, where both a and b correspond to the

positions of nucleotide residues shown in SEQ ID NO:54, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

When tested against K562 cell lines, supernatants removed from cells containing this gene activated the ISRE (interferon-sensitive responsive element). Thus, it is likely that this gene activates kidney and endothelial cells through the Jaks-STAT signal transduction pathway. ISRE is also a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and hematopoetic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the disorders relating to hemopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:143 as residues: Met-1 to His-6, Cys-29 to Ser-49, Pro-72 to Gly-77.

The tissue distribution combined with the biological activity in stimulating the interferon-sensitive responsive element indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages,

including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 583 of SEQ ID NO:55, b is an integer of 15 to 597, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

The translation product of this gene was shown to have homology to the human thromboxane A2 receptor which is known to be a potent stimulator of platelet aggregation (See Genebank Accession No. P21731).This gene maps to chromosome 19, and therefore, may be used as a marker in linkage analysis for chromosome 19.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neutropenia and neutrophia, and other immunological and

hematopoetic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the diseases relating to hemopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 759 of SEQ ID NO:56, b is an integer of 15 to 773, where both a and b correspond to the

positions of nucleotide residues shown in SEQ ID NO:56, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 47

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions:neutropenia and other hemopoetic or immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the disorders relating to hemopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and hemopoetic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:145 as residues: Pro-14 to Pro-28.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or

proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 719 of SEQ ID NO:57, b is an integer of 15 to 733, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neutropenia, and other hemopoetic and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the diseases relating to hemopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:146 as residues: Pro-23 to Tyr-28.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene

product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 517 of SEQ ID NO:58, b is an integer of 15 to 531, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

The translation product of this gene was shown to have homology to the human cathepsin E which is thought to play a role in modulation of the immune system (See Genebank Accession No.P14091). This gene maps to chromosome 1, and therefore, may be used as a marker in linkage analysis for chromosome 1.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hemopoetic or immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological

probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution combined with its homology to a conserved human cathepsin gene indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 838 of SEQ ID NO:59, b is an integer of 15 to 852, where both a and b correspond to the

positions of nucleotide residues shown in SEQ ID NO:59, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

The translation product of this gene was shown to have homology to the human uridine 5' monophosphate synthase which is known to be involved in purine biosynthesis (See Genebank Accession No. P11172). This gene maps to chromosome 3, and therefore, may be used as a marker in linkage analysis for chromosome 3.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hemopoetic or immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hemopoetic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or

immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 666 of SEQ ID NO:60, b is an integer of 15 to 680, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 51

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neurotropenia, and other hemopoetic or immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hemopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:149 as residues: Gln-73 to Gln-82.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of

hematopoetic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:61 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 880 of SEQ ID NO:61, b is an integer of 15 to 894, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:61, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 52

The translation product of this gene was shown to have homology to the enhancer-trap-locus-1 of Mus musculus (See Genebank Accession No. gi|50866) which is thought to be involved in gene regulation pathways during mouse development, particularly in the regulation of homeotic genes. As such, it can be suggested that the protein product of this gene would play a similiar role in humans. One embodiment of this gene comprises polypeptides of the following amino acid sequence:
VKPDPPRAPGENEDSSVPETPDNERKASISYFKNQRGIQYIDLSSDSEDVVSP
N
CSNTVQEKTFNKDTVIIVSEPSEDEESQGLPTMARRNDDISELEDLSELEDLK
DAKLQTLKELFPQRSDN DLLKVIFIGYCSCNDDKISPAFSAIVSSG (SEQ ID NO:216), KDAKLQTLKELFPQRSD (SEQ ID NO:217), KDTVIIVSEPSEDEES

(SEQ ID NO:218), and/or EDSSVPETPDNERKAS (SEQ ID NO:219). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions:hemopoietic or immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:150 as residues: Lys-38 to Gln-46.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, considering the homology to a conserved homeobox protein, would suggest that the protein product of this gene is useful in the detection, treatment, and/or prevention of developmental disorders, particularly those involving the immune system (e.g. immunodeficiencies secondary

to congentital defects or loss of immune organs). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:62 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 677 of SEQ ID NO:62, b is an integer of 15 to 691, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:62, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 53

When tested against PC12 cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response gene 1) pathway. Thus, it is likely that this gene activates sensory neuron cells through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jaks-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

This gene is expressed primarily in human B cell lymphoma and neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune disorders, particularly of B cell related diseases, and disorders related to hemopoiesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, hemopoietic, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the

standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:151 as residues: Met-1 to Asp-12.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. Considering the expression in B-cell lymphomas, this gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, lymphomas, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:63 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 877 of SEQ ID NO:63, b is an integer of 15 to 891, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:63, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neutropenia, and other hemopoietic or immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hemopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:152 as residues: Ser-32 to Cys-37.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:64 and may have been publicly available prior

to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 944 of SEQ ID NO:64, b is an integer of 15 to 958, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:64, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

When tested against PC12 cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response gene 1) pathway. Thus, it is likely that this gene activates sensory neuron cells through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jaks-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions:neutropenia, and other immune or hemopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hemopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, hemopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene

product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:65 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 788 of SEQ ID NO:65, b is an integer of 15 to 802, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:65, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 56

This gene is expressed primarily in fetal liver.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, hepatoblastoma, hepatitis, liver metabolic diseases, and conditions that are attributable to the differentiation of hepatocyte progenitor cells. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the liver, expression of this gene at significantly higher or lower levels may be routinely

detected in certain tissues (e.g.hepatic, developing, or cancerous and wounded tissues) or bodily fluids (e.g.bile, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:154 as residues: His-27 to Arg-34.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various would-healing models and/or tissue trauma. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:66 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1078 of SEQ ID NO:66, b is an integer of 15 to 1092, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:66, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 57

This gene is expressed primarily in IL-1 and LPS induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neutropenia, and other immune or hemopoietic disorders, particularly bacterial infections. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above

tissues or cells, particularly of the diseases relating to hemopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:67 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 720 of SEQ ID NO:67, b is an integer of 15 to 734, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:67, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

This gene is expressed primarily in IL-1 and LPS induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neutropenia, and other hemopoietic or immune disorders, particularly bacterial infections. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the disorders relating to hemopoietic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences,

are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:68 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 692 of SEQ ID NO:68, b is an integer of 15 to 706, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:68, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 59

This gene is expressed primarily in IL-1 and LSP induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, infection, inflammation, in addition to disorders of the immune or hemopoietic systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions.

Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:69 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 422 of SEQ ID NO:69, b is an integer of 15 to 436, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:69, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 60

This gene is expressed primarily in IL-1 and LSP treated neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, bacterial infections, inflammation, in addition to disorders of the hemopoietic or immune systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the

standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:70 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 707 of SEQ ID NO:70, b is an integer of 15 to 721, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:70, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

This gene is expressed primarily in IL-1 and LSP treated neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, bacterial infection, inflammation, in addition to disorders of the hemopoietic or immune systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:159 as residues: Glu-36 to Lys-46.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:71 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present

invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 779 of SEQ ID NO:71, b is an integer of 15 to 793, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:71, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 62

This gene is expressed primarily in IL-1 and LSP treated neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, bacterial infection, inflammation, in addition to immune or hemopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:160 as residues: Gly-18 to Lys-29, Pro-45 to Gly-51, Pro-53 to Lys-58, Pro-72 to Gly-79, Pro-88 to Leu-108, Ala-124 to Ser-134, Ser-138 to Lys-148.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid

arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:72 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 747 of SEQ ID NO:72, b is an integer of 15 to 761, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:72, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 63

This gene is expressed primarily in IL-1 and LSP treated neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune or hemopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:161 as residues: Asp-6 to Glu-15, Pro-76 to Ser-87.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:73 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 659 of SEQ ID NO:73, b is an integer of 15 to 673, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:73, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

This gene is expressed primarily in IL-1 and LSP treated neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune or hemopoietic disorders. Similarly, polypeptides and

antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:74 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 569 of SEQ ID NO:74, b is an integer of 15 to 583, where both a and b correspond to the

positions of nucleotide residues shown in SEQ ID NO:74, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 65

This gene maps to chromosome 19, and therefore, may be used as a marker in linkage analysis for chromosome 19.

This gene is expressed primarily in activated neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, Inflammatory bowel disease, chronic neutropenia (Kostmann's syndrome), chemotherapy induced neutropenia, AIDS, and other immunodeficiency disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hemopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:163 as residues: Gly-17 to Gly-23.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or

immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:75 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 787 of SEQ ID NO:75, b is an integer of 15 to 801, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:75, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 66

This gene is expressed primarily in activated neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, chronic and acute neutropenia, inflammatory bowel disease, neutrophil related multiple organ failure, and other immune or hemopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and hemopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:164 as residues: Met-35 to Glu-51.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:76 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 968 of SEQ ID NO:76, b is an integer of 15 to 982, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:76, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

This gene is expressed primarily in activated neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, acute and chronic neutropenia, inflammatory bowel disease,

neutrophil-related multiple organ failure, and other immune or hemopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and hemopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:165 as residues: Asp-21 to His-26, Leu-31 to His-39, Arg-64 to Thr-70.

The tissue distribution of this gene specifically in neutrophils indicates a possible role in the treatment and/or detection of disease states in which either a lack or excess of neutrophils plays a role in the pathophysiology of the disease state. Targetting this protein could provide a mechanism to inhibit the role of neutrophils in Inflammatory bowel disease and neutrophil releted multiple organ failure. The protein encoded by this gene could be important in the treatment of neutropenia, such as the chronic neutropenic Kostmann's syndronme, AIDS related neutropenia, chemotherapy induced neutropenia, in addition to juvenile periodontis and other states which are caused by decreased neutrophil chemotaxis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:77 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 987 of SEQ ID NO:77, b is an integer of 15 to 1001, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:77, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

This gene is expressed primarily in activated neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, neutropenia, inflammatory bowel disease, neutrophil related multiple organ failure, and other immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune or hemopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and hemopoietic, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:166 as residues: Ile-26 to Ala-34, Thr-81 to Asp-88.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences,

are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:78 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 734 of SEQ ID NO:78, b is an integer of 15 to 748, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:78, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 69

This gene is expressed primarily in adipocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, obesity, and diabetes. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.endocrine, metabolic, or cancerous and wounded tissues) or bodily fluids (e.g. serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:167 as residues: Ser-26 to Lys-36.

The tissue distribution predominantly in adipose tissue, indicates a role in the treatment and/or detection of adipofibrosarcoma, adiponecrosis, obesity and diabetes. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:79 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome.

Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 572 of SEQ ID NO:79, b is an integer of 15 to 586, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:79, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

This gene is expressed primarily in kidney.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, kidney diseases, particularly nephritis and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal and urogenital systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.urogenital, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in kidney indicates that this gene or gene product could be used in the treatment and/or detection of kidney diseases including renal failure, nephritus, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilms Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:80 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide

sequence described by the general formula of a-b, where a is any integer between 1 to 532 of SEQ ID NO:80, b is an integer of 15 to 546, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:80, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 71

This gene is expressed primarily in T-cells and hepatocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune and hepatic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hepatic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, hepatic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and

committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, considering the expression in hepatocytes indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various would-healing models and/or tissue trauma. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:81 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 694 of SEQ ID NO:81, b is an integer of 15 to 708, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:81, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 72

The translation product of this gene was shown to have homology to the human KIAA0213 which is thought to be a serine/threonine protein kinase which may implicate this gene as playing an integral role in signal transduction, particularly in cell cycle regulation (See Genebank Accession No. P25390). When tested against K562 cell lines, supernatants removed from cells containing this gene activated the ISRE (interferon-sensitive responsive element) pathway. Thus, it is likely that this gene activates kidney cells through the Jak-Stat signal transduction pathway. ISRE is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in rhabdomyosarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, rhabdomyosarcoma, and other cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the muscular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.proliferating, muscle, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:170 as residues: Ser-24 to Ala-30.

The tissue distribution in rhabdomyosarcoma tissue combined with its homology to a putative cell cycle modulating protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of cancer and other proliferative disorders, particularly of muscle tissue. Expression within tumor tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division. Alternatively, considering its expression in muscle tissue may suggest indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various muscle disorders, such as muscular dystrophy, cardiomyopathy, fibroids, myomas, and rhabdomyosarcomas. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:82 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 810 of SEQ ID NO:82, b is an integer of

15 to 824, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:82, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 73

When tested against fibroblast cell lines, supernatants removed from cells containing this gene activated the EGR1 (early growth response gene 1 pathway. Thus, it is likely that this gene activates cells through the EGR1 signal transduction pathway. EGR1 is a separate signal transduction pathway from Jaks-STAT, genes containing the EGR1 promoter are induced in various tissues and cell types upon activation, leading the cells to undergo differentiation and proliferation.

This gene is expressed primarily in T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune disorders and disease states. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hemopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well

as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:83 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 775 of SEQ ID NO:83, b is an integer of 15 to 789, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:83, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 74

The translation product of this gene was shown to have homology to the human zinc finger protein 7 which is thought to a play a role as a transcriptional modulator (See Genebank Accesion No. P17097). This gene maps to the chromosome X, and therefore, may be used as a marker in linkage analysis for the chromosome X.

This gene is expressed primarily in T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, immune disorders and disease states. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hemopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having

such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:172 as residues: Glu-4 to Arg-12, Glu-63 to Arg-69.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:84 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 797 of SEQ ID NO:84, b is an integer of 15 to 811, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:84, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 75

This gene is expressed primarily in anergic T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, leukemias, lymphomas, auto-immunities, immunodeficiencies, immuno-supressive conditions and hematopoeitic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoeitic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, hematopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies (e.g. AIDS), immuno-supressive conditions (transplantation) and hematopoeitic disorders. In addition this gene product may be applicable in conditions of general microbial infection, inflammation or cancer. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:85 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1056 of SEQ ID NO:85, b is an integer of 15 to 1070, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:85, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 76

This gene maps to chromosome 11, and therefore, may be used as a marker in linkage analysis for chromosome 11.

This gene is expressed primarily in T-cells (resting and anergic).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, leukemias, lymphomas, auto-immunities, immunodeficiencies, immuno-supressive conditions and hematopoeitic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoeitic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, hematopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.Preferred epitopes include those comprising a sequence shown in SEQ ID NO:174 as residues: Thr-25 to Asp-38.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies (e.g. AIDS), immuno-supressive conditions (transplantation) and hematopoeitic disorders. In addition this gene product may be applicable in conditions of general microbial infection, inflammation or cancer. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:86 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 713 of SEQ ID NO:86, b is an integer of 15 to 727, where both a and b

correspond to the positions of nucleotide residues shown in SEQ ID NO:86, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 77

This gene maps to chromosome 8, and therefore, may be used as a marker in linkage analysis for chromosome 8.

This gene is expressed primarily in anergic T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, leukemias, lymphomas, auto-immunities, immunodeficiencies, immuno-supressive conditions and hematopoeitic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoeitic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, hematopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.Preferred epitopes include those comprising a sequence shown in SEQ ID NO:175 as residues: Glu-8 to Lys-17, Val-42 to Trp-51.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies (e.g. AIDS), immuno-supressive conditions (transplantation) and hematopoeitic disorders. In addition this gene product may be applicable in conditions of general microbial infection, inflammation or cancer. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:87 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded

from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 676 of SEQ ID NO:87, b is an integer of 15 to 690, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:87, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 78

This gene is expressed primarily in anergic T-cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, leukemias, lymphomas, auto-immunities, immunodeficiencies, immuno-supressive conditions and hematopoeitic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoeitic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g. blood cells, and immune, hematopoietic, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of immune disorders including: leukemias, lymphomas, auto-immunities, immunodeficiencies (e.g. AIDS), immuno-supressive conditions (transplantation) and hematopoeitic disorders. In addition this gene product may be applicable in conditions of general microbial infection, inflammation or cancer. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:88 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded

from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 882 of SEQ ID NO:88, b is an integer of 15 to 896, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:88, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 79

The translation product of this gene was shown to have homology to the human clathrin light chain B which is the major protein for the polyhedral coat of clathrin coated pits and vesicles (See Genebank Accession No. P09497).

This gene is expressed primarily in the spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, meningitis, spina bifida, spinal tumors and neoplasms as well as other developmental and neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the spinal cord and central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, or cancerous and wounded tissues) or bodily fluids (e.g.serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in spinal cord combined with the homology to human clathrin indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of trauma, meningitis, spina bifida, spinal tumors and neoplasms as well as other developmental and neurodegenerative conditions of the spinal cord and central nervous system, particularly those neural disorders involving cell-cell signalling. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:89 and may have been publicly available prior to conception of the present invention. Preferably, such related

polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 843 of SEQ ID NO:89, b is an integer of 15 to 857, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:89, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 80

This gene is expressed primarily in spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, meningitis, spina bifida, spinal tumors and neoplasms as well as other developmental and neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the spinal cord and central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, or cancerous and wounded tissues) or bodily fluids (e.g. serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of trauma, meningitis, spina bifida, spinal tumors and neoplasms as well as other developmental and neurodegenerative conditions of the spinal cord and central nervous system, such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

related to SEQ ID NO:90 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 547 of SEQ ID NO:90, b is an integer of 15 to 561, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:90, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 81

This gene is expressed primarily in spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, meningitis, spina bifida, spinal tumors and neoplasms as well as other developmental and neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the spinal cord and central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, or cancerous and wounded tissues) or bodily fluids (e.g. serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:179 as residues: Met-1 to Arg-6, Ser-98 to Met-104.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of trauma, meningitis, spina bifida, spinal tumors and neoplasms as well as other developmental and neurodegenerative conditions of the spinal cord and central nervous system, such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns,

balance, and preception. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:91 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 641 of SEQ ID NO:91, b is an integer of 15 to 655, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:91, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 82

This gene is expressed primarily in spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, meningitis, spina bifida, spinal tumors and neoplasms as well as other developmental and neurodegenerative conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the spinal cord and central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.neural, or cancerous and wounded tissues) or bodily fluids (e.g. serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:180 as residues: Asn-9 to Tyr-14, Ala-30 to Val-39.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of trauma, meningitis, spina bifida, spinal tumors and neoplasms as well as other developmental and neurodegenerative conditions of the spinal cord and central

nervous system, such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:92 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 834 of SEQ ID NO:92, b is an integer of 15 to 848, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:92, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 83

This gene is expressed primarily in fibrosarcoma, and tonsils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions, which include, but are not limited to, fibrosarcoma, tosilitis, and other muscle or immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and musculoskeletal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues (e.g.immune, muscle, or cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of

immune system disorders. Expression of this gene product in tonsils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tumors and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Alternatively, the expression in fibrosarcoma indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection, treatment, and/or prevention of various muscle disorders, such as muscular dystrophy, cardiomyopathy, fibroids, myomas, and rhabdomyosarcomas. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:93 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 598 of SEQ ID NO:93, b is an integer of 15 to 612, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:93, and where the b is greater than or equal to a + 14.

TABLE

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
1	HTECE94	209145 07/17/97	Uni-ZAP XR	11	1761	952	1761	1072	1072	99	1	30	31	65
2	HTWAH05	209145 07/17/97	Lambda ZAP II	12	1519	909	1519	995	995	100	1	39	40	71
3	HAQAN31	209145 07/17/97	Uni-ZAP XR	13	1071	483	1071	503	503	101	1	26	27	50
4	HAUAQ39	209145 07/17/97	Uni-ZAP XR	14	955	65	955	157	157	102	1	21	22	220
5	HBNAU27	209145 07/17/97	Uni-ZAP XR	15	1508	847	1508	942	942	103	1	39	40	113
6	HSIDD28	209148 07/17/97	Uni-ZAP XR	16	2006	1225	2006	1256	1256	104	1	25	26	131

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
7	HCABR41	209145 07/17/97	Uni-ZAP XR	17	545	1	538	15	15	105	1	27	28	87
8	HCUAQ30	209145 07/17/97	ZAP Express	18	602	1	602	61	61	106	1	27	28	63
9	HE2AF21	209145 07/17/97	Uni-ZAP XR	19	587	1	587	237	237	107	1	25	26	40
10	HE2DC87	209145 07/17/97	Uni-ZAP XR	20	644	1	644	308	308	108	1	30	31	48
11	HE2OO64	209145 07/17/97	Uni-ZAP XR	21	1257	142	823	823	823	109	1	18	19	38
12	HE2PO86	209145 07/17/97	Uni-ZAP XR	22	541	39	541	151	151	110	1	32	33	44
13	HE8EV15	209145 07/17/97	Uni-ZAP XR	23	567	1	567	64	64	111	1	22	23	38
14	HEPCE18	209145 07/17/97	Uni-ZAP XR	24	586	1	586	62	62	112	1	21	22	24

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
15	HFCBD73	209145 07/17/97	Uni-ZAP XR	25	1510	892	1510	1134	1134	113	1	12	13	49
16	HKIYA46	209145 07/17/97	pBluescript	26	535	1	535	138	138	114	1	36	37	53
17	HLHSA86	209145 07/17/97	pBluescript	27	1273	763	1273	156	156	115	1	49	50	268
18	HNGJM08	209145 07/17/97	Uni-ZAP XR	28	780	1	780	329	329	116	1	24	25	37
19	HSDSB09	209145 07/17/97	pBluescript	29	819	1	819	22	22	117	1	17	18	121
20	HTGBV53	209145 07/17/97	Uni-ZAP XR	30	608	1	608	136	136	118	1	22	23	33
21	H2CAA57	209145 07/17/97	pBluescript SK-	31	1217	395	1217	282	282	119	1	54	55	178
22	HADFV30	209145 07/17/97	pSport1	32	765	1	765	222	222	120	1	24	25	45

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
23	HAIBO71	209145 07/17/97	Uni-ZAP XR	33	752	172	752	325	325	121	1	28	29	66
24	HAPAT76	209145 07/17/97	Uni-ZAP XR	34	2265	1399	2192	380	380	122	1	30	31	337
24	HAPAT76	209145 07/17/97	Uni-ZAP XR	95	2264	1398	2191	1480	1480	183	1	40	41	65
25	HLHEB47	209147 07/17/97	Uni-ZAP XR	35	643	1	643	37	37	123	1	19	20	68
26	HLHEF54	209147 07/17/97	Uni-ZAP XR	36	1302	437	1302	589	589	124	1	41	42	76
27	HLHFM06	209147 07/17/97	Uni-ZAP XR	37	1005	203	1005	406	406	125	1	17	18	21
28	HLMIG41	209147 07/17/97	Lambda ZAP II	38	608	1	608	154	154	126	1	19	20	36
29	HLMMJ78	209147 07/17/97	Lambda ZAP II	39	925	1	925	59	59	127	1	14	15	61

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
30	HLQBQ85	209147 07/17/97	Lambda ZAP II	40	1219	429	1219	479	479	128	1	34	35	45
31	HLQBR11	209147 07/17/97	Lambda ZAP II	41	1724	1062	1712	1254	1254	129	1	31	32	55
32	HLWBZ56	209147 07/17/97	pCMVSport 3.0	42	798	1	798	262	262	130	1	37	38	62
33	HLYBI18	209147 07/17/97	pSport1	43	693	1	693	28	28	131	1	25	26	34
34	HMAJL22	209147 07/17/97	Uni-ZAP XR	44	1358	342	1358	564	564	132	1	21	22	32
35	HMCAR20	209147 07/17/97	Uni-ZAP XR	45	965	47	958	327	327	133	1	29	30	117
36	HMCAV55	209147 07/17/97	Uni-ZAP XR	46	791	1	791	118	118	134	1	22	23	145
37	HMEFS61	209147 07/17/97	Lambda ZAP II	47	770	1	770	141	141	135	1	23	24	44

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
38	HMEJY78	209147 07/17/97	Lambda ZAP II	48	875	1	875	155	155	136	1	18	19	133
39	HMMAD08	209147 07/17/97	pSport1	49	614	1	614	263	263	137	1	24	25	25
40	HMWFY10	209147 07/17/97	Uni-Zap XR	50	556	1	556		129	138	1			18
41	HMWHH16	209147 07/17/97	Uni-Zap XR	51	1003	1	1003	43	43	139	1	32	33	92
42	HMWID22	209147 07/17/97	Uni-Zap XR	52	886	188	886	558	558	140	1	23	24	40
42	HMWID22	209147 07/17/97	Uni-Zap XR	97	886	188	886	234	234	185	1	20	21	23
43	HNFFC27	209147 07/17/97	Uni-ZAP XR	53	564	1	564	74	74	141	1	19	20	82
44	HNFFC39	209147 07/17/97	Uni-ZAP XR	54	933	1	933	144	144	142	1	35	36	75

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
45	HNGAM20	209147 07/17/97	Uni-ZAP XR	55	597	1	597	266	266	143	1	24	25	92
45	HNGAM20	209147 07/17/97	Uni-ZAP XR	98	597	1	597	232	232	186	1			41
46	HNGDS13	209147 07/17/97	Uni-ZAP XR	56	773	1	773	30	30	144	1	17	18	22
47	HNGDS53	209147 07/17/97	Uni-ZAP XR	57	733	1	733	105	105	145	1	39	40	39
48	HNGDU92	209147 07/17/97	Uni-ZAP XR	58	531	21	531	269	269	146	1	27	28	34
49	HNGED06	209147 07/17/97	Uni-ZAP XR	59	852	1	852	241	241	147	1	25	26	33
50	HNGEW13	209147 07/17/97	Uni-ZAP XR	60	680	1	680	29	29	148	1	30	31	50
51	HNGEY51	209147 07/17/97	Uni-ZAP XR	61	894	1	894	467	467	149	1	51	52	83

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
52	HNGEZ47	209147 07/17/97	Uni-ZAP XR	62	691	1	691	118	118	150	1	46	47	46
53	HNGFQ33	209147 07/17/97	Uni-ZAP XR	63	891	76	891	136	136	151	1	40	41	48
54	HNGFU38	209147 07/17/97	Uni-ZAP XR	64	958	1	958	139	139	152	1	31	32	76
55	HNGIC13	209147 07/17/97	Uni-ZAP XR	65	802	1	802	120	120	153	1	32	33	34
56	HSKXE22	209148 07/17/97	pBluescript	66	1092	202	1092	267	267	154	1	16	17	64
57	HNHBE49	209147 07/17/97	Uni-ZAP XR	67	734	1	734	99	99	155	1	22	23	84
58	HNHBI47	209147 07/17/97	Uni-ZAP XR	68	706	1	706	209	209	156	1			2
59	HNHEC59	209148 07/17/97	Uni-ZAP XR	69	436	1	436	174	174	157	1	23	24	42

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
60	HNHEC63	209148 07/17/97	Uni-ZAP XR	70	721	30	640	124	124	158	1	31	32	31
61	HNHEI54	209148 07/17/97	Uni-ZAP XR	71	793	1	793	184	184	159	1	29	30	46
62	HNHER77	209148 07/17/97	Uni-ZAP XR	72	761	1	761	26	26	160	1	15	16	158
63	HNHES40	209148 07/17/97	Uni-ZAP XR	73	673	1	673	47	47	161	1	30	31	89
64	HNHEV43	209148 07/17/97	Uni-ZAP XR	74	583	1	583	26	26	162	1	61	62	64
65	HNHFL46	209148 07/17/97	Uni-ZAP XR	75	801	1	801	157	157	163	1	39	40	145
66	HNHFP80	209148 07/17/97	Uni-ZAP XR	76	982	1	982	29	29	164	1	28	29	68
67	HNHFS63	209148 07/17/97	Uni-ZAP XR	77	1001	1	1001	688	688	165	1	55	56	89

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
68	HNHGC56	209148 07/17/97	Uni-ZAP XR	78	748	1	748	219	219	166	1	52	53	88
69	HOUCZ78	209148 07/17/97	Uni-ZAP XR	79	586	1	586	143	143	167	1	26	27	36
70	HRAAL86	209148 07/17/97	pCMVSport 3.0	80	546	1	546	148	148	168	1	29	30	33
71	HRDEC77	209148 07/17/97	Uni-ZAP XR	81	708	1	708	539	539	169	1	30	31	36
72	HRDEL61	209148 07/17/97	Uni-ZAP XR	82	824	1	824	159	159	170	1	46	47	57
73	HSAUC38	209148 07/17/97	Uni-ZAP XR	83	789	1	789	178	178	171	1	26	27	44
74	HSAUF49	209148 07/17/97	Uni-ZAP XR	84	811	1	811	23	23	172	1	34	35	95
75	HSAUK57	209148 07/17/97	Uni-ZAP XR	85	1070	1	1070	327	327	173	1	26	27	48

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
76	HSAUL82	209148 07/17/97	Uni-ZAP XR	86	727	1	727	140	140	174	1	25	26	49
77	HSAXI90	209148 07/17/97	Uni-ZAP XR	87	690	1	690	197	197	175	1	44	45	60
78	HSAXN46	209148 07/17/97	Uni-ZAP XR	88	896	1	896		358	176	1	15	16	39
79	HSDGW43	209148 07/17/97	Uni-ZAP XR	89	857	1	857	81	81	177	1	22	23	52
80	HSDJM31	209148 07/17/97	Uni-ZAP XR	90	561	1	561	351	351	178	1	25	26	40
81	HSDJR23	209148 07/17/97	Uni-ZAP XR	91	655	1	655	46	46	179	1	46	47	104
82	HSDMA90	209148 07/17/97	Lambda ZAP II	92	848	114	848	191	191	180	1	31	32	39
83	HSFAM73	209148 07/17/97	Uni-ZAP XR	93	612	1	612	147	147	181	1	13	14	24

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly , the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic

methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a

combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification , such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence

shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown inTable 1, the ORF (open reading frame), or any fragement specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the presence invention can be determined conventionally using known computer programs. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result

of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identiy are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the lenght of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is becuase the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignement of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence

except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is becuase the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are

considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as E. coli).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be

deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions

where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or

the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including

monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the

polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1

and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers,

since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In

this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the

proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect

interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes

Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox , hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteocarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue

regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stoke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit

(antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors),or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity , and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, caricadic rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of

positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type

Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

Vector Used to Construct Library	Corresponding Deposited Plasmid
Lambda Zap	pBluescript (pBS)
Uni-Zap XR	pBluescript (pBS)
Zap Express	pBK
lafmid BA	plafmid BA
pSport1	pSport1
pCMVSport 2.0	pCMVSport 2.0
pCMVSport 3.0	pCMVSport 3.0
pCR®2.1	pCR®2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain

DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ^{32}P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM $MgCl_2$, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is

used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P^{32} using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on

either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

Example 5: Bacterial Expression of a Polypeptide

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 ($O.D.^{600}$) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high

affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

Example 6: Purification of a Polypeptide from an Inclusion Body

The following alternative method can be used to purify a polypeptide expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfuidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem

columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 μg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring

signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five μg of a plasmid containing the polynucleotide is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One μg of BaculoGold™ virus DNA and 5 μg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 μl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture

and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 μl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μCi of ^{35}S-methionine and 5 μCi ^{35}S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No.209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five µg of the expression plasmid pC6 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of metothrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 µM, 2 µM, 5 µM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 µM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:
GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGCC
CAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCCAAAACC
CAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGTGGT
GGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACG
GCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAAC
AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTG
AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAACCCCC
ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGT
GTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCT
GACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGTGGAGTGGGA
GAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGG
ACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCA
GGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC
ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGC
GACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera

containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2 x 10^5 cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in

Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl2 (anhyd); 0.00130 mg/L $CuSO_4$-$5H_2O$; 0.050 mg/L of $Fe(NO_3)_3$-$9H_2O$; 0.417 mg/L of $FeSO_4$-$7H_2O$; 311.80 mg/L of Kcl; 28.64 mg/L of $MgCl_2$; 48.84 mg/L of $MgSO_4$; 6995.50 mg/L of NaCl; 2400.0 mg/L of $NaHCO_3$; 62.50 mg/L of NaH_2PO_4-H_20; 71.02 mg/L of Na_2HPO4; .4320 mg/L of $ZnSO_4$-$7H_2O$; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H_20; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H_20; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H_20; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalainine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tryrosine-2Na-$2H_20$; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B_{12}; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine;

0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

Example 12: Construction of GAS Reporter Construct

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proxial region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

Ligand	tyk2	JAKs Jak1	Jak2	Jak3	STATS	GAS(elements) or ISRE
IFN family						
IFN-a/B	+	+	-	-	1,2,3	ISRE
IFN-g		+	+	-	1	GAS (IRF1>Lys6>IFP)
Il-10	+	?	?	-	1,3	
gp130 family						
IL-6 (Pleiotrohic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
Il-11(Pleiotrohic)	?	+	?	?	1,3	
OnM(Pleiotrohic)	?	+	+	?	1,3	
LIF(Pleiotrohic)	?	+	+	?	1,3	
CNTF(Pleiotrohic)	-/+	+	+	?	1,3	
G-CSF(Pleiotrohic)	?	+	?	?	1,3	
IL-12(Pleiotrohic)	+	-	+	+	1,3	
g-C family						
IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP >>Ly6)(IgH)
IL-7 (lymphocytes)	-	+	-	+	5	GAS
IL-9 (lymphocytes)	-	+	-	+	5	GAS
IL-13 (lymphocyte)	-	+	?	?	6	GAS
IL-15	?	+	?	+	5	GAS
gp140 family						
IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)
IL-5 (myeloid)	-	-	+	-	5	GAS
GM-CSF (myeloid)	-	-	+	-	5	GAS
Growth hormone family						
GH	?	-	+	-	5	
PRL	?	+/-	+	-	1,3,5	
EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)
Receptor Tyrosine Kinases						
EGF	?	+	+	-	1,3	GAS (IRF1)
PDGF	?	+	+	-	1,3	
CSF-1	?	+	+	-	1,3	GAS (not IRF1)

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:
5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCG
AAATGATTTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:
5':<u>CTCGAG</u>ATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAATG
ATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCC
CTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC
CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGC
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTT
TGCAAA<u>AAGCTT</u>:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, Il-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml genticin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies)

with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final concentration of 10^7 cells/ml. Then add 1ml of 1 x 10^7 cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest $2x10e^7$ U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM $Na_2HPO_4.7H_2O$, 1 mM $MgCl_2$, and 675 uM $CaCl_2$. Incubate at 37^oC for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37^oC for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting $1x10^8$ cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of $5x10^5$ cells/ml. Plate 200 ul cells per well in the 96-well plate (or $1x10^5$ cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37^oC for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon

activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS

(Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as 5×10^5 cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to 1×10^5 cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

NF-κB (Nuclear Factor κB) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF-κB regulates the expression of genes involved in immune cell activation, control of apoptosis (NF-κB appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κB is retained in the cytoplasm with I-κB (Inhibitor κB). However, upon stimulation, I- κB is phosphorylated and degraded, causing NF- κB to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κB include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF-κB promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF-kB would be useful in treating diseases. For example, inhibitors of NF-κB could be used to treat those diseases related to the acute or chronic activation of NF-kB, such as rheumatoid arthritis.

To construct a vector containing the NF-κB promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF-κB binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:
5':GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGAC
TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:
5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene) Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCA
TCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACT
AATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTC
CAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT:
3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF-κB/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF-κB/SV40/SEAP cassette is removed from the above NF-κB/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF-κB/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF-κB/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described

in Example 13. As a positive control, exogenous TNF alpha (0.1,1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 µl of 2.5x dilution buffer into Optiplates containing 35 µl of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 µl Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 µl Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25

28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO_2 incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 ul of 12 ug/ml fluo-3 is added to each well. The plate is

incubated at 37°C in a CO_2 incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10^6 cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10^6 cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca^{++} concentration.

Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating

tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford,MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford,MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na3VO4, 2 mM Na4P2O7 and a cocktail of protease inhibitors (# 1836170) obtained from Boeheringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and

PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg_{2+} (5mM ATP/50mM $MgCl_2$), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM $MgCl_2$, 5 mM $MnCl_2$, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30^{o}C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37^{o}C for 20 min. This allows the streptavadin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phospotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37^{o}C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR

products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Manheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10.

The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 μg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 μg/kg/hour to about 50 μg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally,

intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other

disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 116, line N/A

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet

Name of depositary institution
American Type Culture Collection

Address of depositary institution (*including postal code and country*)
10801 University Boulevard
Manassas, Virginia 20110-2209
United States of America

Date of deposit July 17, 1997	Accession Number 209147

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only	For International Bureau use only
☒ This sheet was received with the international application	☐ This sheet was received by the International Bureau on
Authorized officer Perry Hackley International Division 703-305-8517 perryhackley@uspto.gov	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis)*

A. The indications made below relate to the microorganism referred to in the description on page 113 , line N/A

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet

Name of depositary institution
American Type Culture Collection

Address of depositary institution *(including postal code and country)*

10801 University Boulevard
Manassas, Virginia 20110-2209
United States of America

Date of deposit July 17, 1997	Accession Number 209145

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only	For International Bureau use only
☒ This sheet was received with the international application	☐ This sheet was received by the International Bureau on
Authorized officer Perry Hackley International Division 703-305-6517 [illegible]	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 113, line N/A.

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet

Name of depository institution
American Type Culture Collection

Address of depositary institution (*including postal code and country*)

10801 University Boulevard
Manassas, Virginia 20110-2209
United States of America

Date of deposit July 17, 1997	Accession Number 209148

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only	For International Bureau use only
☒ This sheet was received with the international application	☐ This sheet was received by the International Bureau on
Authorized officer Perry Hackley International Division 703-305-6517 perry.hackley@uspto.gov	Authorized officer

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;

(b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;

(c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;

(d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;

(e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;

(f) a polynucleotide which is a variant of SEQ ID NO:X;

(g) a polynucleotide which is an allelic variant of SEQ ID NO:X;

(h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;

(i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.

3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

9. A recombinant host cell produced by the method of claim 8.

10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;

(c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(g) a variant of SEQ ID NO:Y;

(h) an allelic variant of SEQ ID NO:Y; or

(i) a species homologue of the SEQ ID NO:Y.

12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.

14. A recombinant host cell that expresses the isolated polypeptide of claim 11.

15. A method of making an isolated polypeptide comprising:

(a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and

(b) recovering said polypeptide.

16. The polypeptide produced by claim 15.

17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.

18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

(a) contacting the polypeptide of claim 11 with a binding partner; and

(b) determining whether the binding partner effects an activity of the polypeptide.

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

22. A method of identifying an activity in a biological assay, wherein the method comprises:

(a) expressing SEQ ID NO:X in a cell;

(b) isolating the supernatant;

(c) detecting an activity in a biological assay; and

(d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 22.

<110> Rosen
et al.
Human Genome Sciences, Inc.

<120> 83 Human Secreted Proteins

<130>

<140> Unassigned
<141>

<150>
<151>

<160> 219

<170> PatentIn Ver. 2.0

<210> 1
<211> 733
<212> DNA
<213> Homo sapiens

<400> 1
```
gggatccgga gcccaaatct tctgacaaaa ctcacacatg cccaccgtgc ccagcacctg    60
aattcgaggg tgcaccgtca gtcttcctct tccccccaaa acccaaggac accctcatga   120
tctcccggac tcctgaggtc acatgcgtgg tggtggacgt aagccacgaa gaccctgagg   180
tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg   240
aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg caccaggact   300
ggctgaatgg caaggagtac aagtgcaagg tctccaacaa agccctccca acccccatcg   360
agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc   420
catcccggga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct   480
atccaagcga catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga   540
ccacgcctcc cgtgctggac tccgacggct ccttcttcct ctacagcaag ctcaccgtgg   600
acaagagcag gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc   660
acaaccacta cacgcagaag agcctctccc tgtctccggg taaatgagtg cgacggccgc   720-
gactctagag gat                                                      733
```

<210> 2
<211> 5
<212> PRT
<213> Homo sapiens

<220>
<221> Site
<222> (3)
<223> Xaa equals any of the twenty naturally ocurring L-amino acids

<400> 2
```
Trp Ser Xaa Trp Ser
  1               5
```

<210> 3
<211> 86
<212> DNA
<213> Homo sapiens

<400> 3
gcgcctcgag atttccccga aatctagatt tccccgaaat gatttccccg aaatgatttc 60
cccgaaatat ctgccatctc aattag 86

<210> 4
<211> 27
<212> DNA
<213> Homo sapiens

<400> 4
gcggcaagct tttgcaaag cctaggc 27

<210> 5
<211> 271
<212> DNA
<213> Homo sapiens

<400> 5
ctcgagattt ccccgaaatc tagatttccc cgaaatgatt tccccgaaat gatttccccg 60
aaatatctgc catctcaatt agtcagcaac catagtcccg cccctaactc cgcccatccc 120
gcccctaact ccgcccagtt ccgcccattc tccgccccat ggctgactaa ttttttttat 180
ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 240
ttttggaggc ctaggctttt gcaaaaagct t 271

<210> 6
<211> 32
<212> DNA
<213> Homo sapiens

<400> 6
gcgctcgagg gatgacagcg atagaacccc gg 32

<210> 7
<211> 31
<212> DNA
<213> Homo sapiens

<400> 7
gcgaagcttc gcgactcccc ggatccgcct c 31

<210> 8
<211> 12
<212> DNA
<213> Homo sapiens

<400> 8
ggggactttc cc 12

<210> 9
<211> 73
<212> DNA
<213> Homo sapiens

<400> 9
gcggcctcga ggggactttc ccggggactt tccggggact ttccgggact ttccatcctg 60
ccatctcaat tag 73

<210> 10
<211> 256
<212> DNA
<213> Homo sapiens

<400> 10
ctcgagggga ctttcccggg gactttccgg ggactttccg ggactttcca tctgccatct 60
caattagtca gcaaccatag tcccgcccct aactccgccc atcccgcccc taactccgcc 120
cagttccgcc cattctccgc cccatggctg actaattttt tttatttatg cagaggccga 180
ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggcttttttg gaggcctagg 240
cttttgcaaa aagctt 256

<210> 11
<211> 1761
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (20)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (24)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (30)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (32)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (546)

<223> n equals a,t,g, or c

<400> 11
ttcgaaaacg attagtgaan gtanaagtan cngcagtacg gtcggattcc cgggtcgacc 60
cacgcgtccg ggaagatgag gaggtcgctg agagctggga agaggcggca gacagcgggc 120
aggaaatcca aatctcctcc caaagtgccc attgtgattc aggacgatag ccttcccgcg 180
gggccccctc cacagatccg catcctcaag aggcccacca gcaacggtgt ggtcagcagc 240
cccaactcca ccagcaggcc caccсttcca gtcaagtccc tagcacagcg agaggccgag 300
tacgccgagg ccсggaagcg gatcctgggc agcgccagcc ccgaggagga gcaggagaaa 360
cccatcctcg acaggccaac caggatctcc caacccgaag acagcaggca gcccaataat 420
gtgatcagac agcctttggg tcctgatggg tctcaaggct tcaaacagcg cagataaatg 480
caggcaagaa aagatgccgc cgttgctgcc gtcaccgcct cctgggtcgt ccgccacggg 540
ttgcantgcc gtggcagaca gctggacttg agcagaggga acgacctgac ttacttgcac 600
tgtgatcccc cttgctccgc ccactgtgac cttgaacccc atgcactgtg acctcccccc 660
ttctccccct tcccactgtg attggcacat cgacaagggc tgtcccaagt caatggaaag 720
ggaaagggtg ggggttaggg gaaggttggg gggacccagc aaggactcag agagtcagac 780
agtgccactt ggccacttgg ggtaaagcca gtgccagcaa taacagttta tcatgctcat 840
taatttggga tttcaaaaca caaatgaaaa ctcacaccca cccacccсca agtgcatgtc 900
tccatcactt aaaaagtaag ttccatttga aaatatcctt tctttttttt ttcttcctat 960
ttttgtttgt ttatacaaat atctgatttg caagaaaaag tgcatgggag gggttttagt 1020
ggtttaatga atttttaatt aagaaagggt agtttggtag tctacttaaa aatgtttctg 1080
ggaaattcac tagaaacatt aaccaatagg attttggtga gcttagcttc tgtattccta 1140
ctgccgccca gaaaaggggc agggctctgc agccgccagg acagacgagc accccatgcc 1200
tatacctccc tccccgagct aagtcccagg gcatctgggc cttgcctgga gactgggcta 1260
gctctgtagg ctcggagctg ggagggtgcc caaccccacc tctagtattt tgggagatag 1320
ggaaagtgaa ccgacttccc cttcccatac ccctcagggt ggttccctac cagccaggct 1380
tactacttct agaagaaagc agagtgccag ggagtgagat tgcatccctg ggcttagaag 1440
tgacggagag aagacttgtt tagtattttg ccatcagcac aaggaaaacc aggagagagt 1500
ctgcctccag gactctgagc cttctgcctc gtatgttcag aaggtggata ggtcttccca 1560
ctccagcatg gcttgaactc ttaggggtct gcagtgctcc atctccattg gtggccccag 1620
ctcagtaact atacctggta catttcctgt gtgcaatcag taccttgaag gcagaacatt 1680
ctgaataaag ttggaaaaar aamaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1740
aaaaaaaaaa aaaaaactcg a 1761

<210> 12
<211> 1519
<212> DNA
<213> Homo sapiens

<400> 12
tcgacccacg cgtccggcgc gcaccgccgc ccgggcagcc gctggctcca gctcacgaaa 60
cagccccggg cgccgcgccg ctctgagtcc agcctcctac tgagaacagt ccctcccttg 120
tgcgggtcgc aggctagccg caggttcggc cacgtcaaat ccatttcta aaaaagcagg 180
gagcagagct ctctcttcgc cgccgacgca gaaaggagct ggggaggaaa aagctgctgc 240
cttttgcgct ggagattcgt gggcaaggct tctcattttc ccaggctgct tcccctcccg 300
ggtgaggagc gtcctgagac taaggaaaga gcctggaaaa tggagcagac ctggacgaga 360
gattattttg cagaggatga tggggagatg gtacccagaa cgagtcacac agcagctttt 420
cttagtgaca ctaaagatcg aggccctcca gtgcagtcac agatctggag aagtggtgaa 480
aaggtcccgt ttgtgcagac atattccttg agagcatttg agaaaccccc tcaggtacag 540
acccaggctc ttcgagactt tgagaagcac ctcaatgacc tgaagaagga gaacttcagc 600
ctcaagctgc kcatctactt cctggaggag cgcatgcaac agaagtatga ggccagccgg 660
gaggacatct acaagcggaa cactgagctg aaggttgaag tggagagctt gaaacgagaa 720
ctccaggaca agaaacagca tctggataaa acatgggctg atgtggagaa tctcaacagt 780
cagaatgaag ctgagctccg acgccagttt gaggagcgac acagkgagac ggagcatgtt 840

```
tatgagctct tggagaataa gatscagctt ctgcaggagg aatccaggct agcaaagaat    900
gaagctgcgc ggatggcagc tctggtggaa gcagagaagg agtgtaacct ggagctctca    960
gagaaactga agggagtcac caaaaactgg gaagatgtac caggacacca ggtcaagccc   1020
gaccaataca ctgaggccct ggcccagagg gacaagtagg tgccttcggt gctctttttg   1080
tcgcttgtct tttgcccatt ctcaaggcat acagcagctg tcctgttccc tttcaaggac   1140
tgacagtagg agcttcacta tttctaagac tttatgggcc cacaaccgaa gacattcttt   1200
tcagggttga attttcagtg gtatccatta tgaaaactca cttcatggat tcagtgggca   1260
aatagcggca agcaagagac atggattcac ttattcggca aacatttact gggcatgcca   1320
catgccagat accgggctaa gtatctggca tgtgttacag aaacaaaaga cctaaatctt   1380
gtcaccaaga aacatgttac atgattttaa taagttccct gatagaagag catggggtgc   1440
tctggggaaa tattggaggg tcatccattc cacattaaaa gagcaagttg tctgcaaaaa   1500
aaaaaaaaaa aaaactcga                                                1519
```

<210> 13
<211> 1071
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (158)
<223> n equals a,t,g, or c

<400> 13

```
gccgcgcacc agttccctga ggtacggttc tccatggtgc acaagcggat caacctggcg     60
gaggacgtgc tggmctggga gcacgagcgc ttcgccatcc gccgactgcc cgccttcacg    120
ctgtcccacc tggagagcca ccgtgacggc cagcgcanag catcatggac gtgcggtccc    180
gggtggattc taagaccctg acccgtaaca cgaggatcat tgcagaggcc ctgactcgag    240
tcatctacaa cctgacagag aaggggacac ccccagacat gccggtgttc acagagcaga    300
tgatccagca ggagcagctg gactcggtga tggactggct caccaaccag ccgcgggccg    360
gcagctggtg gacaaggaca gcaccttcct cagcacgctg gagcaccamc tgagcmgcta    420
cctgaaggac gtgaagcagc accacgtcaa ggctgacaag cgggacccag agtttgtctt    480
ctacgaccag ctgaagcaag tgatgaatgc gtacagagtc aagccagccg tctttgacct    540
gctcctggcc gttggcattg ctgcctacct cggcatggcc tacgtggctg tccagcactt    600
cagcctcctc tacaagaccg tccagaggct gctcgtgaag gccaagacac agtgacacag    660
ccaccccac agccggagcc cccgccgctc cacagtccct ggggccgagc acgagtgagt    720
ggacactgcc ccgccgcggg cggccctgca gggacagggg ccctctccct ccccggcggt    780
ggttggaaca ctgaattaca gagctttttt ctgttgctct ccgagactgg ggggggattg    840
tttcttcttt tccttgtctt tgaacttcct tggaggagag cttgggagac gtcccggggc    900
caggctacgg acttgcggac gagcccccca gtcctgggag ccggccgccc tcggtctggt    960
gtaagcacac atgcacgatt aaagaggaga cgccgggacc cccaaaaaaa aaaaaaaaaa   1020
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaactcga a            1071
```

<210> 14
<211> 955
<212> DNA
<213> Homo sapiens

<400> 14

```
gggcggcggc gcgcgggtgc gatcgcggag ctgtgaggcg caggcagggc tctggggcac     60
ctagagaccg gggccggaga cgtggcagcc gccctgcccg ccagaaagtt tcctagaagt    120
ttgctgggcg cggcgcacg actgactggc tggaccatga acgtgttccg aatcctcggc    180
gacctgagcc acctcctggc catgatcttg ctgctgggga agatctggag gtccaagtgc    240
```

tgcaagggca tctctgggaa gagccagatc ctgtttgctc tcgtcttcac caccaggtac 300
ctggacctgt tcaccaactt catctccatc tacaacacag taatgaaggt ggtttttctc 360
ctctgtgcct atgttacagt gtacatgata tatgggaaat tccgtaaaac ttttgacagt 420
gagaatgaca cattccgcct ggagtttctt ctggtcccag tcattggcct ttccttcctt 480
gaaaactaca gtttcactct gctggagatc ctctggactt tctctatcta tctggaatca 540
gtggctatcc tgccccagct cttcatgatc agcaagactg gagaggctga gaccataact 600
actcactacc tgttctttct gggtctgtac cgggcactct acctggctaa ctggatcagg 660
cggtaccaga ctgagaattt ctatgaccaa attgcagtcg tgtctggagt agtacaaacc 720
atcttctact gtgacttctt ctacttgtat gggaccaaag gtaggtcctg ggatgacagc 780
aatgctgaca ctggcctaag gagttactca tccatttaat aagtattcca gcagatacag 840
atgtgaacag tcaagtctct gccatccaca atgcttgtgt tctaatgcaa gaagacaaat 900
attttcaata aagaaacaaa tgccataaaa acaaaaaaaa aaaaaaaaaa ctcga 955

<210> 15
<211> 1508
<212> DNA
<213> Homo sapiens

<400> 15
ctctgaactg cgccatcaac tccagtctcg gcagcagctc cgctcccggc gccacccacc 60
gacaccccca gaaccctctg ggggcctgcc caggggaccc cctgagcccc ccgaccggct 120
tagctgtgat gggagtcgag tgcatttgct ttataagtga gggtagggtg agggaggaca 180
ggccagtagg gggagggaaa gggagagggg caagggcagg ggactcagga agcaggggt 240
ccccatcccc agctgggaag aacatgctat ccaatctcat ctcttgtaaa tacatgtccc 300
cctgtgagtt ctgggctgat ttgggtctct catacctctg ggaaacagac ctttttctct 360
cttactgctt catgtaattt tgtatcacct cttcacaatt tagttcgtac ctggcttgaa 420
gctgctcact gctcacacgc tgcctcctca gcagcctcac tgcatctttc tcttcccatg 480
caacaccctc ttctagttac cacggcaacc cctgcagctc ctctgccttt gtgctctgtt 540
cctgtccagc aggggtctcc caacaagtgc tctttccacc ccaaagggcc tctcctttc 600
tccactgtca taatctcttt ccatcttact tgcccttcta tactttctca catgtggctc 660
cccctgaatt ttgcttcctt tgggagctca ttcttttcgc caaggctcac atgctccttg 720
cctctgctct gtgcactcac gctcagcaca catgcatcct cccctctcct gcgtgtgccc 780
actgaacatg ctcatgtgta cacacgcttt tcccgtatgc tttcttcatg ttcagtcaca 840
tgtgctctcg ggtgccctgc attcacagct acgtgtgccc ctctcatggt catgggtctg 900
cccttgagcg tgtttgggta ggcatgtgca atttgtctag catgctgagt catgtctttc 960
ctatttgcac acgtccatgt ttatccatgt actttccctg tgtaccctcc atgtaccttg 1020
tgtactttct tcccttaaat catggtattc ttctgacaga gccatatgta ccctaccctg 1080
cacattgtta tgcacttttc cccaattcat gtttggtggg gccatccaca ccctctcctt 1140
gtcacagaat ctccatttct gctcagattc cccccatctc cattgcattc atgtactacc 1200
ctcagtctac actcacaatc atcttctccc aagactgctc ccttttgttt tgtgtttttt 1260
tgaggggaat taaggaaaaa taagtggggg caggtttgga gagctgcttc cagtggatag 1320
ttgatgagaa tcctgaccaa aggaaggcac ccttgactgt tgggatagac agatggacct 1380
atggggtggg aggtggtgtc cctttcacac tgtggtgtct cttggggaag gatctccccg 1440
aatctcaata aaccagtgaa cagtgtgact cggaaaaaaa aaaaaaaaaa aaaaaaaaaa 1500
aaactcga 1508

<210> 16
<211> 2006
<212> DNA
<213> Homo sapiens

<220>
<221> SITE

<222> (70)
<223> n equals a,t,g, or c

<400> 16
ggtcacgagg cacttttggt aagccaggga tcgtgtgggt ggtacggagt tgccacattt 60
cggcaaaggn aactatgtag ctgatcttgg agccatggtg gtaacaggtc tttgagggaa 120
tcctatggct gtggtcagca aacaagtaaa tatggaactg gccaagatca agcaaaaatg 180
cctactttat gaagccaacg gacaagctga cactgtcaag gttcctaaag agaaagatga 240
aatggtagag caagagttta accggttgct agaagctaca tcttacctta gtcatcaact 300
agacttcaat gtcctcaata ataagcctgt gtcccttggc caggcattgg aagttgtcat 360
tcagttacaa gagaagcatg tcaaagatga gcagattgaa cattggaaga agatagtgaa 420
aactcaggaa gaattgaaag aacttcttaa taagatggta aatttgaaag agaaaattaa 480
agaactccat cagcaataca aagaagcatc tgaagtaaag ccacccagag atattactgc 540
cgagttctta gtgaaaagca aacacaggga tctgaccgcc ctatgcaagg aatatgatga 600
attagctgaa acacaaggaa agctagaaga aaaacttcag gagttggaag cgaatccccc 660
aagtgatgta tatctctcat caagagacag acaaatactt gattggcatt ttgcaaatct 720
tgaatttgct aatgccacac ctctctcaac tctctccctt aagcactggg atcaggatga 780
tgactttgag ttcactggca gccacctgac agtaaggaat ggctactcgt gtgtgcctgt 840
ggctttagca gaaggcctag acattaaact gaatacagca gtgcgacagg ttcgctacac 900
ggcttcagga tgtgaagtga tagctgtgaa tacccgctcc acgagtcaaa cctttattta 960
taaatgcgac gcakttctct gtacccttcc cctgggtgtg ctgaagcagc agccaccagc 1020
cgttcagttt gtgccacctc tccctgagtg gaaaacatct gcagtccaaa ggatgggatt 1080
tggcaacctt aacaaggtgg tgttgtgttt tgatcgggtg ttctgggatc caagtgtcaa 1140
tttgttcggg catgttggca gtacgactgc cagcargggt gagctcttcc tcttctggaa 1200
cctctataaa gctccaatac tgttggcact agtggcagga gaagctgctg gtatcatgga 1260
aaacataagt racgatgtga ttgttggccg atgcctggcc attctcaaag ggatttttgg 1320
tagcagtgca gtacctcagc ccaaagaaac tgtggtgtct cgttggcgtg ctgatcccta 1380
tgttgctgca ggatcatctg gaaatgacta tgatttaatg gctcagccaa tcactcctgg 1440
cccctcgatt ccaggtgccc cacagccgat tccacgactc ttctttgcgg gagaacatac 1500
gatccgtaac tacccagcca cagtgcatgg tgctctgctg agtgggctgc gagaagcggg 1560
aagaattgca gaccagtttt tgggggccat gtatacgctg cctcgccagg ccacaccagg 1620
tgttcctgca cagcagtccc caagcatgtg agacagatgc attctaaggg aagaggccca 1680
tgtgcctgtt tctgccatgt aaggaaggct cttctagcaa tactagatcc cactgagaaa 1740
atccaccctg gcatctgggc tcctgatcag ctgatggagc tcctgatttg acaaaggagc 1800
ttgcctcctt tgaatgacct agagcacagg gaggaacttg tccattagtt tggaattgtg 1860
ttcttcgtaa agactgaggc aagcaagtgc tgtgaaataa catcatctta gtcccttggt 1920
gtgtggggtt tgtttttttt ttttatattt tgagaataaa acttcatata aaattgaaaa 1980
aaaaaaaaaa aaaaaaaaaa actcga 2006

<210> 17
<211> 545
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (530)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (540)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (545)
<223> n equals a,t,g, or c

<400> 17
gaattcggca cgagatggaa aacaccttct ttgtcttcct ggtgtcagcc ctgctgctgg 60
ccgtgatcta cctcaacatc caggtggtgc ggggccagcg caaggtcatc tgcctgctca 120
aggagcagat cagcaatgag ggtgaggaca aaatcttctt aatcaacaag cttcactcca 180
tctacgagag gaaggagagg gaggagagga gcagggttgg gacaaccgag gaggctgcgg 240
caccccctgc cctgctcaca gatgaacagg atgcctaggg ggacggcgat gggcctcacg 300
ggccsgccca gcaccctgag accacactgt tgcctcccag tgaccctgct gggacaccag 360
gacaaggaag acagtttcgc ctctcgaaag ccgcagctgc gcctaggctg gagctggaag 420
ggtgggtgaa tccggcttgg gcatccccaa tgaactctgc cctgcctggg actctattta 480
ttctgattaa aggggttttg caaatgaaaa aaaaaaaaaa aaaaaaaccn cgggggggn 540
ccggn 545

<210> 18
<211> 602
<212> DNA
<213> Homo sapiens

<400> 18
gaattcggca cgagtgccct gggttccgat tgatacccac tcttgttgtt caaaagagag 60
atgagtgctg cttctttttg gccccggcct gtggcaagta tctcagtctt catactgctg 120
ggaagctctg taaccaccag caagaccaga agtggggtga tcagcagtgc aggaaagccc 180
atttgggtgc agtccccgca cctagcccctt ttggaagtgc ttctccaaaa gggaattgtg 240
ccggaaaagt agggattgaa accaaacagc cacatcctgc catcaggatg ctctttatgg 300
ccccactgac caagaaatca cagcttctgt actcagtgat gactgcttga cttcagttga 360
ggaaaacaat gaagttctgt agccaggcgt ggtggcagat gtctgtaatc ccagctactc 420
gggaggctga ggcaagagaa ttgcttaaac cccgggaggt ggaggttgca gtgagccgag 480
atggcgacac tgactccagc ctgggtgcca gagcgagact ctttgtctca aaaaaaaaaa 540
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaagaaa aaaacggcac 600
ga 602

<210> 19
<211> 587
<212> DNA
<213> Homo sapiens

<400> 19
ggctgcagga attcggcacg agtaaggcta tataaacggg aagttaagta ttaatagaat 60
tcagtgtgac taacaaggga tatcgagttc ttcagacctg tggtatacat aataagcttg 120
aagaatttgt gcacaaaagt cttaactgtt ttgcagcctt ggttgtggtt agatgctgta 180
gttttctaag ctactacatg ttgtaaccag cttggtagat ttaacctagt aataagatgc 240
ttagtttaac tgttagttta aagtcagttt ctatagcggc acagtcttta tttttggacc 300
ttcactttcc aatccagatg acacttgtcc ataaagaaat tgctaaactt gagacctaaa 360
aacaaaacaa aaaacaaaaa aactacagac aagtaacctt taaaattatt tcgcttgatg 420
gaaatttacc ggaaggcttt aaccaattca gtttgcttag actcataaag aaaattatga 480
taatgtctag gtaaacttca gcaaaacatt tttttgtgaa attatactat agtacataaa 540
attgaaatat tgctgattta atgtaaaaaa aaaaaaaaaa aactcga 587

<210> 20
<211> 644
<212> DNA
<213> Homo sapiens

<400> 20
aattcggcac gagagattac atccctcctg tgttcacacc ccaccatgac tttacaatga 60
tcttacagtg caagagtctt tacaagaccc tacaatacat gatcttcctc gcagcatttt 120
acctctcccc tatttcttaa tacacttctt cttcactctg ggccagctac actaaccttc 180
ttgttgtttc ttgtatactt atcagacaca ttcttccctc aagctatttg tatttgctat 240
tccttcaaat aaccttatag cttgttctta acatccttca aaccattgct cagatgtaca 300
cttccacatg actctttacc ttaacactaa caaaaataaa ccgtctgccc tgtactctct 360
ctttttctgc tttatttcca ccccatatac ttatggcctt caaatatgct ataaatgttt 420
ttttatttat atttttgtta tctgtctcta ctaaaaatac aaaaattagc tgggtgtgct 480
ggctggcacc tgtaatccca gctacttggg aggcttaggc agaagaatca cttgaacctg 540
ggaggcagag gttgcagtga gctgagatcg cgccactgta ctccagcctg ggcaacagga 600
gtgagatccg tctcaacaaa aaaaaaaaaa aaaaaaaaac tcga 644

<210> 21
<211> 1257
<212> DNA
<213> Homo sapiens

<400> 21
aggaaatgct ttgggatgag tctattcttg gattttgaat gtttagtttt gtttacccaa 60
ggttgaattg aaaaaaaaaa acagtcaata tggatttaga aaaaggaaca cctgatgaag 120
aaaaggagag gtagatacag tcagtgtcac ttcaggacac ttaggttttt tttgtataaa 180
aatttaaatt gaattaaaag aaggaaaaaa aaagcccaaa cttaacctct gagaaagaac 240
ataagaactc aaggagaaca taagagaaaa ggaaacctgt tacagaaaag acaagaatct 300
gtgttttgga atgagtctat tcttgggtat tgaacttttta gttttgtttg cccaaggatt 360
aattgaggaa atcagctaag aaaatggact ttagacaaaa gcaagaggat cagatgaaga 420
aaaggagagg tagatacagt cagtgtcact tcaggaaagc tatttaaaaa aacttgaaat 480
ttaactgaaa gaagaaacaa caacaaaaaa gcctaaacct agcctctgaa caacactaac 540
atgagaacac aagaacttaa gagaaaaaga aacctactca agaaaagaca gaagagacag 600
tgatttggga tgagtctact ctaggatttt caacttttta gttttgttcc ttcaaagttg 660
aaggaaaaaa agtttggttt tataaaattc atgttattgt aatttttcta ggtggatggc 720
tattttaatc tctaaaaaag ccaagtgaag taaaagtatt cagtatgcct tttcctcaag 780
ttactttcct tcattttctt aaaaaaraaa aaaaattatt aaatgtttct cacatatctc 840
acatataatg taatttccct aaatgaagtt gtctctactt ctgctcatca aattgctgtg 900
atagtgaatt atttattcat gggagataat ttattttaaa ggacagaatt accaagcgtt 960
acaaaatcag ttctttcctt ggttttgtgt tagtgttggt ggtattttat tgttgttttt 1020
ctgtgtttat gtgtctcagc tttctccaag gaatatgtat gaaataactt aaactgattt 1080
tttctttgtt aaatctaatt tgcagtgtat ttttgcattt tctagttctg aaagtggaaa 1140
atgaaacagt ctataataaa cttagatgat atatagtttt aaaacggtct caaaaagtac 1200
tgatataagg tcagtctata ttctggaaat gtttatatta aagtgtttta atttcta 1257

<210> 22
<211> 541
<212> DNA
<213> Homo sapiens

<220>
<221> SITE

<222> (2)
<223> n equals a,t,g, or c

<400> 22
gntccaattt cattttaaaa gatgtagaaa gaagaatcaa gcatcaatta attataaagc 60
ctaaagcaaa gttagatttg ggggttattc agccaaaatt accgttttag accagaatga 120
atagactaca ctgataaaat gtactggata atgccacatc ctatatggtg ttatagaaat 180
agtgcaagga aagtacattt gtttgcctgt cttttcattt tgtacattct tcccattctg 240
tattcttgta caaaagatct cattgaaaat ttaaagtcat cataatttgt tgccataaat 300
atgtaagtgt caataccaaa atgtctgagt aacttcttaa atccctgttc tagcaaacta 360
atattggttc atgtgcttgt gtatatgtaa atcttaaatt atgtgaacta ttaaatagac 420
cctactgtac tgtgctttgg acatttgaat taatgtaaat atatgtaatc tgtgacttga 480
tattttgttt tatttggcta tttaaaaaca taaatctaaa aaaaaaaaaa aaaaaactcg 540
a 541

<210> 23
<211> 567
<212> DNA
<213> Homo sapiens

<400> 23
gaattcggca cgagctggat tcctttttg tctctaacaa aatatctaaa gaaaaccgaa 60
aacatgctcc gtataaaaag ttgtctctta ttattttta ttttttttcc atttaacatt 120
aaagactctc aagtacctgc caattatatt gccacatttt ctaggaaatg cagcttttag 180
caattctttg ttgattcaaa tgaaatcaac ctagctcagc taatattaat tgattagatt 240
gagaataaag tcctaatacc aaaggctgac caagagaaaa tgcttgaaat cagatgttga 300
ctgattcagg ccggttctat cagtttgggc aagttgctag ggagtggaca ggaagcttga 360
ggacatcaca aaagaatcca taaaggaccc atgatgcatt gagagacaga tacataagaa 420
tggctgggca tagtagaaca gatctggtat cattacagta aatctccatt atatggagtt 480
atctagaaac attatcttcc ttgctggctg aagaaacata gtacccctcc aactaccctc 540
aaacaaaaaa aaaaaaaaaa actcgta 567

<210> 24
<211> 586
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (28)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (550)
<223> n equals a,t,g, or c

<400> 24

```
nccgctctta gaactagtga tcccccgngc tgcaggctgg ggggctccgg ttccctgagg    60
gatgagcctt cagcctccct ttgtaatgct gctcctctcc actgcccagc accatgagtt   120
gggtgcagac acctagaagg agagacttct tggaacgctc atcccccgct atacctcccc   180
ttcctcctgc atctcccctt ctttccttcc ccctcaggag agagaaaact tagtgcttcc   240
agcccttctt ggagccttca tggtccaggg gtaggggccc cactggcctg agcatgccat   300
tttgagggga gggtagttgt gcctacttat cccctggcag aggggatgcc aggaccatgg   360
acatgaggct tgcccatccc tgccaactta cacagcctgt accactgtcc ccccttcctt   420
ggctactttg acatgtgcct gctcctggca tttcaataaa acccggcttg ggtctgaaaa   480
aaaaaaaaaa aaaaaaactc gaggggggc cggtamccaa ttcgccctat artgaatcgt   540
attaaaattn aatgggcggt cgttttacaa agtcgtgact ggggaa                  586


<210> 25
<211> 1510
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (39)
<223> n equals a,t,g, or c

<400> 25
aggcgacgtg ggccagatct tctcctgccg cctgcaggnc accaacaact tcttcggcgc    60
cgggcagaac aagcggccgc caagctgggc magatcggcc ggagcaagcg ggttgttatt   120
gaagatgata ggattgatga cgtgctgaaa aatatgaccg acaaggcacc tcctggtgtc   180
taactccccc aaagacaatg agttaaggga gagaataaga acggcggtaa cagttattgg   240
caaaaagcat gaaaagagaa agcactttga aatttattac tagcttgcta cccacgatga   300
aatcaacaac ctgtatctgg tatcaggccg ggagacagat gaggcgagag gaggaggagg   360
aggaggagaa ggctctgggc tcctctgcaa aaataaaaat aaaaaaataa ataaaatttt   420
aaaaataata aaaattcact atatacacat ataaagaaat aaaaagaagt ctcagttgca   480
gctatttgtc aaaattaata tccatttctt tttatatacg gtgaatattg cgcaattata   540
gatctggatt ttgaaccact taatgaagcg gcaacaccag gtgttttgag gtgttggcat   600
tcttcgctga tttggctgtt cccaatgttt acattattta atcttgcaaa aatggttctg   660
tgcacttgga tgtgaaatgc tgtccagttt tatttttttt atgttgttat ccttggatgt   720
acaaaaaatt cagaaaatga tctctgtaga tattctgttt tattttggtc atctttagaa   780
gttatcagga atgtgtttaa aacaagaaga gaacttttct aaggaatgat acatagaaaa   840
gattttattt taaaatgagt tgtaaagctt gtgtttcttt gttgctgcaa gctatctgcc   900
caagttaatg caaatggaca catttttttat gtcagaaaaa cacacacaca cacacacaca   960
cacacacaca cacacacaca cgaaaaacaa agaaaaaaat gcttgagctt tttctaactt  1020
ccccttgcag tctgttgtgt gagcagcctg tttatttctc taatattatg tcagtttatt  1080
ctctttaatg gactgtaaaa aaatgtaatc acaagagtgc caaatatctt gaaatgccaa  1140
aaggcatttt agtttctttt ctctgtgctc tgagtccacg tacaggaatg cttggagtgt  1200
cttttctgtt atttataggg attctcttaa ggcacaccag ctgcctgttt tgcatggtat  1260
ttgcaaaaat gcctcttgcg tgaggaaatc ttttaccatt ttttgtttgc aactttggac  1320
ctcaagaggt ttcccttccc ttccccgttc cctcttttct taattcaata ttctgtatgt  1380
tgcaccttga accagcacac agggctattt ctccaatgta caataaaaga attgttcctg  1440
tgtctcaaaa aaaaaaaaaa aaaaaaactc gaggggggc ccgtacccaa tcgcctratg  1500
atcgtatagc                                                         1510


<210> 26
<211> 535
<212> DNA
<213> Homo sapiens
```

<220>
<221> SITE
<222> (529)
<223> n equals a,t,g, or c

<400> 26
tccacgcgtt ggcggccgct ctagaactag tggatccccc ggsctggsag aattcggcas 60
grgccccggt caggccctgc ccagagagct cctggttcct gaactgagct gcctctaccg 120
tggtgggctg ggcaggcatg tgccccccta gtcagagggc accaacccac ctactctgcc 180
cctgggtgga tcctgggccg gtcgtgttag ggttgtccct ctgggtgctg gctggtggga 240
tgggkgaggg tggggagcag ctcccagcac ccctgctgtg tggttcatct ttttttagg 300
cccctgcctg tctgcccatc tgcccctcac ccaccctagg ctctgcccac cgcctggacc 360
tgcccacccc tgaaagactg gcccctggct ccccgcccct cggtctccac gtggtgtatg 420
gatctgtggt cattgtccct ctgcagaata aagattgctc aggcctgcct ggaaaaaaa 480
aaaaaaaaaa aaaaaaaact cgagggggg cccgtaccca atcgcctgng atgat 535

<210> 27
<211> 1273
<212> DNA
<213> Homo sapiens

<400> 27
tcccccgggc ctgcaggaat tcggcagacc cgrctgtaaa gatggcggct tcctagtgag 60
tcggcggctg atttagaagg aggttcaggc tacggtgagc cgaagggagg attctggctt 120
cccctgtccg tgttccatct agccacacag gagccatgga agtggcagag cccagcagcc 180
ccactgaaga ggaggaggag gaagaggagc actcggcaga gcctcggccc cgcactcgct 240
ccaatcctga aggggctgag gaccgggcag taggggcaca ggccagcgtg ggcagccgca 300
gcgagggtga gggtgaggcc gccagtgctg atgatgggag cctcaacact tcaggagccg 360
gccctaagtc ctggcaggtg cccccgccag cccctgaggt ccaaattcgg acaccaaggg 420
tcaactgtcc agagaaagtg attatctgcc tggacctgtc agaggaaatg tcactgccaa 480
agctggagtc gttcaacggc tccaaaacca acgccctcaa tgtctcccag aagatgattg 540
agatgttcgt gcggacaaaa cacaagatcg acaaaagcca cgagtttgca ctggtggtgg 600
tgaacgatga cacggcctgg ctgtctggcc tgacctccga cccccgcgag ctctgtagct 660
gcctctatga tctggagacg gcctcctgtt ccaccttcaa tctggaagga cttttcagcc 720
tcatccagca gaaaactgag cttccggtca cagagaacgt gcagacgatt cccccgccat 780
atgtggtccg caccatcctt gtctacagcc gtccaccttg ccagccccag ttctccttga 840
cggagcccat gaagaaaatg ttccagtgcc catatttctt ctttgacgtt gtttacatcc 900
acaatggcac tgaggagaag gaggaggagg atgaagccat tgaggttgag gccactgtct 960
gaaccatccc tgtacatctg caccttcttg tgcaaggaag tccttggcct aaagccttgg 1020
ttctcaaact gggttccttg ggacctccgg ggtggggggg ttccaggagg cacgtagggt 1080
accttgcagg gtcctaggag ggaaacccag gattccagga gggatcccag gaactgtggg 1140
cacccatttt ctgtgtctcc cagcccattt ccactcctag tttgtcatgg ataatttttg 1200
ttcttccctg tgtgattttt gccatcaaaa taaaatttg agactcgtta aaaaaaaaaa 1260
aaaaaaaact cga 1273

<210> 28
<211> 780
<212> DNA
<213> Homo sapiens

<400> 28
gaattcggca cgagcggacg ggacagggga cctggctgtt gagctgtcca agaccgaccc 60

ggcgagcctg gagacaggcc aggacagtga ggatgactca ggcgagccag aggactgggt 120
cccggaccct gtggatgccg atccaggtta gcttgcccca catggtccct ttcaccagtg 180
tgggggcgc ttcagccagg ccgtccgctc tggactcatc ccttgtcaca gagcttggct 240
gtgtcaggtg tcgctagtgt ctcagaggct tgagggtgtc aaaggtcagg gcagtgctcc 300
gccacctgcc agcctgggca ggccggtgat gggctgtttt cctctgtggc tggtcactct 360
ggccgtgggg gatgccctgc ctcccaccgc ttgtgagctg tggggtgtcc ctgccccgcc 420
actgcacctg gcagaggagt gaagatgtgg gtgacccctg agctctgtac tgatttgggc 480
ctcaaagtcc cgttagggtg ccagctcctg ctcagagccc atctgtgcct ggccctcgtc 540
tgcgagggct ctgagggctg tggggcccag gcctggcccc tgccctggca cagatggtaa 600
gtcttcaaca catctttagt ctctctcaga ggaaaacccc aggacctttg gcttgtcccc 660
taagcctgca tgggagtccc ggaaggtccc tgtggggtga gaggggctgg ggattgaagc 720
tgttgctcca ggacctgact ccaggtctcc cgggaggctc agggctgctc ccagcctcga 780

<210> 29
<211> 819
<212> DNA
<213> Homo sapiens

<400> 29
gaattcggca cgaggagaat catgggcctc tggctgggca tgctggcctg tgtcttcctg 60
gcaactgctg cctttgttgc ttatactgcc cggctggact ggaagcttgc tgcagaggag 120
gctaagaaac attcaggccg gcagcagcag cagagagcag agagcactgc aaccagacct 180
gggcctgaga aagcagtcct atcttcagtg gctacaggca gttcccctgg cattaccttg 240
acaacgtatt caaggtctga gtgccacgtg gacttcttca ggactccaga ggaggcccac 300
gccctttcag ctcctaccag cagactatca gtgaaacagc tggtcatccg ccgtggggct 360
gctctggggg cggcgtcagc acactgatgg tggggctcac ggtcaggatc ctagccacca 420
ggcactagca aagaagcttg gaaatagaaa gccaggagtg gctgtcccca gtatgcaaac 480
acaccacggt ctgccctgca aaaacaccaa tggggtctag tgcaggtgga cactttgaac 540
cactcctcaa aaaaagaact ttggctgaty ccttgtggtg acactcagag ggtctgaac 600
agacttgaca attctgttct ggtcaagctg gagttttctt ctgtgacttg gactgctcta 660
cagaagacat cagccaactg cacgagtcag agtccaggga ttgtcactat tattaataat 720
gtaaatggct tcaaatggga cactgcagat aammycacaa aaaccactgt tatattaaag 780
attacacatt tcctggaaaa aaaaaaaaaa aaaactcga 819

<210> 30
<211> 608
<212> DNA
<213> Homo sapiens

<400> 30
gaattcggca cgagcttcac tccttgttcc ctctcctgca cacacacaca cacacacaca 60
ccccaggaga ctctacccca actcagccct aacccagccg aacaaccttc agtggctccc 120
cagtgcctga agaatatgat ccaaactttc cctgcctacc tctgtcttcc cctttttac 180
gtgctggacc tggcccttgc ttctgctcct gtactttctc attcagctct cctctgacat 240
gtttcttcc ttccccacaa ggccagcaaa aatatcacct cctccaggaa gtcctccatg 300
accagtgagc wcaccaggaa cttctgcttt tgagccccca gaaagagcca tgctgcagaa 360
gtctccccat gctgcttcta acccaaataa agtacaggag aggagttcag gaaaaagtct 420
agagccaggc acagcagtac acgcctctac tctcagctam tcgggaggct gaggtgggag 480
gattgcttga acccaggagt ctgaggctgc tatgcaytat gatcacacct gtgattagcc 540
actgtactcc agcctgggca acatagcaag aacctgtttc ttaaaaaaa aaaaaaaaa 600
aactacga 608

<210> 31
<211> 1217
<212> DNA
<213> Homo sapiens

<400> 31
cgttacacat gacaccagtg cctttgtttc attgggctgg gctctctgga aggtgtgctg 60
ctgcctgagc tgctggaaaa gcactgacag gtgtttgcta gaaaagcact cctggagctt 120
gccaccagct tggacttcta gggactttcc tctcagccag gaaggatttt gatattcatc 180
agaaatacct ccagaagatt caaggagctg tagaggtgaa gtaagcctgt gaaggaccag 240
catggaatc ctatactctg agcccatctg ccaagcagcc tatgcagaat gactttggac 300
aagtgtggcg gtgggtgaaa gaagacagca gctatgccaa cgttcaagat ggctttaatg 360
gagacacgcc cctgatctgt gcttgcaggc gagggcatgt gagaatcgtt tccttccttt 420
taagaagaaa tgctaatgtc aacctcaaaa accagaaaga gagaacctgc ttgcattatg 480
ctgtgaagaa aaaatttacc ttcattgatt atctactaat tatcctctta atgcctgtyc 540
tgcttattgg gtatttcctc atggtatcaa agacaaagca gaatgaggct cttgtacgaa 600
tgctacttga tgctggtgtc gaagttaatg ctacagattg ttatggctgt accgcattac 660
attatgcctg tgaaatgaaa aaccagtctc ttatccctct gctcttggaa gcccgtgcag 720
accccacaat aaagaataag catggtgaga gctcactgga tattgcacgg agattaaaat 780
tttcccagat tgaattaatg ctaaggaaag cattgtaatc cttgtgacca caccgatgga 840
gatacagaaa aagttaacga ctggattcta tcttcatttt agacttttgg tctgtgggcc 900
atttaacctg gatgccacca ttttatgggg ataatgatgc ttaccatggt taatgttttg 960
gaagagcttt ttatttatag cattgtttac tcagtcaagt tcaccatggc cgtaatcctt 1020
ctaagggaaa cactaaagtt gttgtagtct ccacttcagt cagaaactga tgtttcagct 1080
aggcacagtg gtacatgcct gtaatcccag ctacttggga ggctgaggtg ggaggatcac 1140
ttgaactcag gagtttgaga gcagccaggg caacacagcg agaccctgtc tcaaaaaaaa 1200
aaaaaaaaaa aactcga 1217

<210> 32
<211> 765
<212> DNA
<213> Homo sapiens

<400> 32
ccacgcgtcc ggtgaggtct catgtctgct tatgcggtgg ctcgctgctc agaacaggga 60
accattggag atactcatta ctctttgaag gcttacagtg gaatgaattc aaatacgact 120
tatttgagga attgaagttg actttatgga gctgataaga atcttcttgg agaaaaaaag 180
actggtactt ctgaattaac caaaatcaca gtattctgaa gatgattcta caaagcctgc 240
tgtttctaca aaggctgctg atgatttcta caaagcctgc tgtagtgttg ctgtggcctc 300
tgcttaaaaa agtagaaaac acattgatgc agcatgttca ccccaacctc cctgcctaaa 360
ggctcaggga ccatcttgga agaggaaggc gcgtgagatt gtaagagccg aattaggggg 420
atggagtgtg gagaataagg acacttcatc ttggatgctc acctgccaaa ttgacttctg 480
atgaaagcca gctccagaaa tgtgcctaca gttactactt tcacctaaac cctgccctta 540
gtcaaatcct tctcttcttc taagcaatca acttcaattc cttgtataac ccacagtata 600
aaagggcttt tataccattc tatcctattg catgtaagcc ttgggtctgg gaggtaacag 660
tgtgggattc caccatctca tctccctgcc acccaaacat gcctgctctt ctttaagcaa 720
tattaaatgt ttgtacttca gaaaaaaaaa aaaaaaaggg cggcc 765

<210> 33
<211> 752
<212> DNA
<213> Homo sapiens

<400> 33
actgaacagt ggttaatcct gactctgttt ttgactgaca gttaacagtt acatgaacca 60
ttcatattac agctcttact taaatttgac caagccagga tatatctgtt aggccacatt 120
catttaggga tcatgttttc caaagcaggt ttgggcaaaa ttaatccaca ggactgaaag 180
gtatacatct gtgagttttg ttctcacttc cacctctaat ttgaagaaca ctttaattga 240
cacagaatac atttcacata tttaacctct acaataagtt ctgacacatt ttccatgaaa 300
caaaccatcg ctatattcaa gataatgaac ctatctatca tactcccaaa ttccttctkg 360
catctttgta atttctcact cttccttctc cctctccccg tcccatccca accactgatc 420
tgctcaggca actaccaatc ttctttctgt cactatagat taatttgcat ttttaaagaa 480
atttacatac atggaaccat acatcatcta tgctttgrag tatgactcct gtcactcagt 540
acaattattt tgagattcat ttatgttawt gtatgtatca atagttcatc ccttttattg 600
gtaagtaaca tttttttgta taggtatacc atgatttgtt gatgaacaaa tttacctgtt 660
gatgaacatt tacgttgtta ccaagatttt tgctattgaa aataaagttt ttatgaatat 720
ttatatatat aaaaaaaaaa aaaaaaactc ga 752

<210> 34
<211> 2265
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (300)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (2162)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (2258)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (2265)
<223> n equals a,t,g, or c

<400> 34
aaatttctca acaccacagt cakctwagtc acctactgcc accttcgaaa aacacggaga 60
gcacctaccc agaggagaag gtagatttgg agtaagccgc cgtcgacata attcctctga 120
tggttttttt aacaatggtc ccctacgaac tgcaggagat tcttggcacc agscctccct 180
gttccgccat gattctgtgg actctggtgt ctctaaggga gcatatgctg gaatcacagg 240
gaacccatct ggttggcata gctcttcccg aggtcatgat ggcatgagcc aacgtakggn 300
aggtggcaca gggaaccatc gccattggaa tggcagcttc cactcccgga aagggtgtgc 360
ttttcaggaa aagccaccta tggagattag ggaagaaaag aaagaagaca aggtggaaaa 420
gttgcagttt gaagaggagg actttccttc cttgaatcca gaagctggca aacagcatca 480
gccatgcaga cctattggga caccttctgg agtatgggaa aacccgccta gtgccaagca 540
accctccaag atgctagtta tcaaaaaagt ttccaaagag gatcctgctg ctgccttctc 600
tgctgcattc acctcaccag gatctcacca tgcaaatggg aacaaattgt catccgtggt 660
tccaagtgtc tataagaacc tggttcctaa gcctgtacca cctccttcca agcctaatgc 720
atggaaagct aacaggatgg agcacaagtc aggatccctt tcctctagcc gggagtctgc 780

ttttaccagt ccaatctctg ttaccaaacc agtggtactg gctagtggtg cagctctgag 840
ttctcccaaa gagagtccct ccagcaccac ccctccaatt gagatcagct cctctcgtct 900
gatcaagttg acccgccgaa ccaccgacag gaagagtgag ttcctgaaaa ctctgaagga 960
tgatcggaat ggagacttct cagagaatag agactgtgac aagctggaag atttggagga 1020
caacagcaca cctgaaccaa aggaaaatgg ggaggaaggc tgtcatcaaa atggtcttgc 1080
cctccctgta gtggaagaag gggaggttct ctcacactct ctagaagcag agcacaggtt 1140
attgaaagct atgggttggc aggaatatcc tgaaaatgat gagaattgcc ttcccctcac 1200
agaggatgag ctcaaagagt tccacatgaa gacagagcag ctgagaagaa atggctttgg 1260
aaagaatggc ttcttgcaga gccgcagttc cagtctgttc tcccccttgga gaagcacttg 1320
caaagcagag tttgaggact cagacaccga aaccagtagc agtgaaacat cagatgacga 1380
tgcctggaag taggcatata aatgctcaca gttaaatctg acccagtaaa ctctgtgtgt 1440
ttagggagta tacaaaagaa atcgttcttt tcctttcctt atgttgttga atacttcatt 1500
cacaagggaa ataatcatat cccaaagaga gagcaattgg cttgttttgc ttttgttatt 1560
gttcttccct gttatctgct ttatagagag aagtttgtgt ggtgggacag atttttaaa 1620
cacactcaya cacacacaca catacacacc cagtatatat ggggcgatgc acaggtagga 1680
gctggcagtg cagggaagag gagacactgg tctgcagcaa cagcttctac taccagccct 1740
tggggcactc acccctgtga tcaagcaatc attgtcaatg acaaagtgac tattgaagtt 1800
ataattgtat taaattaatg ctaataattt ggatatttta ttttattttt ggctgctcgg 1860
gtaactttag cccttaacca agcatatgtg ggtttttttg gttgtttttt tttgtttttt 1920
ttttcttttt ccttttgggg tacagctgta aaatatttgg atataggaaa tgttgtgtta 1980
ttcttgcagc cttgatattc agggtggatt gtaaaatata aatttttgtg agatttcaaa 2040
gattaagatt attttgataa cattatttac agatttaaaa gatgtggtta tcacaagtct 2100
cgagggggaa actactgcat aaaataacta acttggaata aatattttgc atcagtttgg 2160
anaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2220
aaaaaaaaaa aaaaaaaaaa aaaaaaaaag ggggggncc ccccn 2265

<210> 35
<211> 643
<212> DNA
<213> Homo sapiens

<400> 35
gaattcggca cgagctgctg tggggccaaa cgcatcatga aggaagctct ccactgggcc 60
cttttcagca tgcaagccac gggccacgtg ctgcttcacc tcctgttacc tgcagcagct 120
cctcgatgcc acagaggacg ggcatccccc caagggcaag gcctcatccc tcatcccgac 180
ctgtctgaag atactgcagt gaaagcccaa gccctagctt tccccagtga aggactagac 240
taggggcccc acgctcaact ggtagtggcc acaagcctgg cagctgtaga gccgctaacc 300
tcccgacacc tccctcacca cacaggaccc tgagtgagga ggaggggctg gaaacctggg 360
rtgggttggc caaaggagaa cctcaggctc ctggcctggc ccagctcctt cctgcccaag 420
gtagcttagc ccatccagac tggtcctgaa gtctgtccct ccattggcat gaagtctgcc 480
cctcagcagt ccggcctcac aggctgtact ttcatggtgc tctctacctt ctggccccca 540
tcccagaaca ttcgtgagtg aattcgcaag catactagca tgtgatatta gggagtttgc 600
aataaattat tgatgctgaa aaaaaaaaaa aaaaaaaact cga 643

<210> 36
<211> 1302
<212> DNA
<213> Homo sapiens

<400> 36
cctccggcgg gagccacctc tccggggcgc atcatcckgc ccsscagtgc tgtccttatt 60
cccagcccag tcaagagcta ccggggctgg ctagtcatgg gggagcccag tagagaggag 120
tataaaatcc agtcctttga tgcagagacc cagcagctgc tgaagacagc actcaaagat 180

```
ccgggtgctg tggacttgga gaaagtggcc aatgtgattg tggaccattc tctgcaggac    240
tgtgtgttca gcaaggaagc aggacgcatg tsctacgcca tcattcaggc agagagtaaa    300
caagcaggcc agagtgtctt ccgacgtgga ctcctcaacc ggctgcagca ggagtaccag    360
gctcgggagc agctgcragc acgctccctg cagggctggg tctgctatgt cacctttatc    420
tgcaacatct ttgactacct gagggtgaac aacatgccca tgatggccct ggtgaaccct    480
gtctatgact gcctcttccg gctggcccag ccagacagtt tgagcaagga ggaggaggtg    540
gactgtttgg tgctgcagct gcaccgggtt ggggagcagc tggagaaaat gaatgggcag    600
cgcatggatg agctctttgt gctgatccgg gatggcttcc tgctcccaac tggcctcagc    660
tccctggccc agctgctgct gctggagatc attgagttcc gggcggccgg ctggaagaca    720
acgccagctg cccacaagta ttactacagc gartctccga ctaggcytcc agatcagggc    780
ttcctcacca gcactggcct ttcttctacc cacctctaaa gctggcagtg gagtctctgc    840
ctcacccaaa gacttttccc ttccagactt tgagtgtctt cccttctaga ctttcccatc    900
tcctggtgag atgtttccca cttatgccgt ggtcctgccc tgagcccctt tccccaccac    960
aacccaccac ggccaggcag agaagggcaa ctcccaagag ccactgcact gtgtaaccat   1020
tagtgcaact actaccttgg tgcctcagtt tacccatctg taaaatgggt aagcatagcc   1080
actggtggga tattttggga tgtcaagggg tggaggcaga gcacaagtca caccagaaac   1140
tgctttttat acattttgta taaggacaac tctggaaaca agcctatttc ctccagccag   1200
tttcactgaa tgctgcacca catgctacac cagttcagcg tgagaatttt ctaataaatc   1260
ttttctgata ctaaaaaaaa aaaaaaaaaa aaaaaaactc ga                      1302
```

<210> 37
<211> 1005
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (488)
<223> n equals a,t,g, or c

<400> 37

```
caggaaacca aatgatgtcc ctgcccgccg cccccgccgg gcggctttcc cccttgtact     60
ggagaagctc gaacacccgg tcacagctct ctttgctatg ggaactggga cacttttta     120
cacgatgttg ccgccgtccc caccctaacc cccacctccc ggccctgagc gtgtgtcgct    180
gccatatttt acacaaaatc atgttgtggg agccctcgtc ccccctcctg cccgctctac    240
cctgacctgg gcttgtcatc tgctggaaca ggcgccatgg ggcctgccag ccctgcctgc    300
caggtcccct agcacctgtc cccctgcctg tctccagtgg gaaggtagcc tggccaggcg    360
gggcctcccc ttcgacgacc aggcctcggt cacaacggac gtgacatgct gcttttttta    420
attttatttt tttatgaaaa gaaccagtgt caatccgcag accctctgtg aagccaggcc    480
ggccgggncg agccagcagc ccctctccct agactcagag gcgccgcggg gaggggtggc    540
cccgccgagg cttcaggggc cccctcccca ccaaagggtt cacctcacac ttgaatgtac    600
aacccacccc actgtcggga aggcctccgt cctcggcccc tgcctcttgc tgctgtcctg    660
tccccgagcc cctgcaggtc cccccccgcc cccccactca agagttagag caggtggctg    720
caggccttgg gcccggaggg aaggccactg ccggccactt ggggcagaca cagacacctc    780
aaggatctgt cacggaaggc gtcctttttc cttgtagcta acgttaggcc tgagtagctc    840
ccctccatcc ttgtagacgc tccagtccct actactgtga cggcatttcc atccctcccc    900
tgcccgggaa gggaccttgc agggacctct ccctccaaaa aaagaaaaaa agaaaaaraa    960
araaaaaawa aaaactccga ggggggcccc ggtacccaat tcgcc                   1005
```

<210> 38
<211> 608
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (73)
<223> n equals a,t,g, or c

<400> 38
ccctcactaa ggggaccaaa agctggagct ccaccgcggt ggcggccgct ctagaactta 60
gtggatcccc cgngctgcag gaattcggca cgagaaacac tgaaattggg cagagtagaa 120
gtcagcattt agtgagacca aataaaatta gtaatgggca agctgctgtt tcctttgctt 180
ttagctccat tttctcccat aaacaaatat attcttcact ttgcaagaga tggagtagaa 240
gaagttttga aatttgtatc ctaaattagc ttcaagtaag tgcctaaaga gacctctttc 300
ccttaaaacc tgttaatcag ttaaaggcgg ggaacactgg tgcctttttt tttttttttt 360
taacttctta accaagggac agtgaagact tttaagttag atctgatttt agaattgcag 420
ttgaggtagt gcctagtgtg tgaatttgag gtcattttct aaactggccg ggcacagtgg 480
ctcatgcctg taatcccagc actttgggag gcccaggtgg gagaatcact tgagtccagg 540
agtttgatac cagcctgagc aacacaggga gaccccatct ctacaaaaaa aaaaaaaaaa 600
aaactcga 608

<210> 39
<211> 925
<212> DNA
<213> Homo sapiens

<400> 39
gaattcggca cgagagaaat tctgccccgcc tctgtaaggc acagtacaca cataagacat 60
gcttgttgta gctgtgatct ttcttcatgg cgctggggct atgaactatc taattgctaa 120
gattctggag gtgcagggcc tcagggaggt gccatgcaca tacaatacaa ggggtatagc 180
gccccctgga ggcaacgttg gatttgaagc agccagtgtg gtggacaggc cttgtgggca 240
gtgaggaggc ctgattctaa ccttggccct gcagagaact ccctggctat ggaactcagg 300
ggcagatctc accgaacttc tctgggcctc aaatccacat ctgtaaaagg aaggggttgt 360
actctcaatg gttctaaaac tccttctctt cctctgtccc ttcctttctg catcttcttg 420
gcagaatcct ttcctcaaat cagtgctccg atagacccct aatacagtaa acagctgaag 480
gtggagagtc tctgttggaa gtgggtcttg gcctaagcct ccctcattga cccatgatga 540
ctgcmcagaa gagatttcca aactcctccc agccctagca gttatgggtc cagggcctca 600
tcactgagaa cagccccaca ccagtttact tttcattttt tagagtgagg gttttactct 660
gtcacctagg aggcggaggt cagaatgagt tgagatcaca ccactgcact tcagtctggg 720
caacagagtg aggctctgtc tttaaaaaaa aaaaaaaaaa aagaaagggc tggctgtggt 780
ggctcactcc tgtaatccca atcccagcac tttgggaggc tgaggcgggc ggatcgcctg 840
aggtcgggaa ttcaagacca gcctgaccaa tatggagaaa ccccgtctct actaaaaata 900
caaaaaaaaa aaaaaaaaaa tccgc 925

<210> 40
<211> 1219
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (3)
<223> n equals a,t,g, or c

<220>

<221> SITE
<222> (19)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (90)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (599)
<223> n equals a,t,g, or c

<400> 40
gcncacgagc cgcggcggna agtcggttcc cgtgacgcgg cgcgccccaa gggccggctc 60
cgttgaggga agggaagccc gcccggtggn ggctggggtc ggctgctggg aggaggtggt 120
gggctggttc ggacgtgggt cgaggctgta gcaggactcc aggaagatgt taccgagtac 180
ttcagtgaat tccttagtgc aggggaacgg agtcttgaat tccagggatg cggcaagaca 240
cacagccgga gcgaaacgct acaaatatct gagaaggctt ttccgctttc ggcaaatgga 300
ctttgaattt gctgcctggc agatgctcta cctgttcaca tccccacaga gagtttacag 360
aaattttcat tatcgaaaac agacgaagga ccagtgggcc agagatgacc ctgctttctt 420
ggtcctgtta agtatctggc tctgtgtgtc cactatagga tttggctttg tgctggacat 480
gggattcttt gagacaataa agcttctcct ttgggttgta ctcatagatt gtgtaggcgt 540
tggtcttctg atagcaactt taatgtggtt catctctaac aagtatttag tgaaacganc 600
agagcagaga ctatgatgtg gaatggggct atgcttttga tgtgcatctc aatgcttttt 660
atccactcct ggtcattttg cattttatcc agcttttttt catcaaccat gttatcctga 720
cagacacatt tattggatat ttagttggaa ataccttatg gttggttgca gttggctatt 780
atatctatgt aactttcctg ggatacagtg cattgccatt tttgaaaaat acagtaattc 840
ttctgtatcc atttgcacct ctgattctgc tctacgggct ttccctggca ctgggatgga 900
acttcaccca tactctctgt tctttctata agtacagagt gaaataaaaa gtgagaagaa 960
gattcaatcg taactgtgtc aacagtattg tgaagtgatc atttcttgta aaacttgtaa 1020
ataaactatc atctttgtag atatcttaaa ggtgtaaagt ttgcaaattt gaagaaatat 1080
atattaacac tgtggtcagg tacattcctt aaaactaatt aaatgtacat ttctataata 1140
aatatttttt aaactaaaaa aaaaaaaaaa aaactcgagg gggggcccgg tacccaattc 1200
gccctatagt gagtcgtat 1219

<210> 41
<211> 1724
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (51)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (371)
<223> n equals a,t,g, or c

<400> 41
agaggggcta tgattcggag ggttctgccg cacggcatgg gccggggcct nttgacccgg 60

aggccaggca cgmgcrgagg aggcttttcw cwsgrattca caggcctcca ctgggctctc 120
agtaggtggg atttgtcagc agttctcatt gcctgagaac atccagctct ccctcttttc 180
agccagcttt ctacggaaca tcaatgagta cctggccgta gtcgatgctc ctcccctgga 240
cctccggttc aacccctcgg gctacctctt gctggcttca gaaaaggatg ctgcagccat 300
ggagagcaac gtgaaagtgc agaggcagga gggagccaaa gtttctctga tgtctcctga 350
tcagcttcgg naacaagttt ccctgggata aacacagagg gagtggcttt ggcgtcttat 420
gggatggagg acgraggttg gtttracccc tggtgtctgc tccaggggct tcggcgaaag 480
gtccagtcct tgggagtcct tttctgccag ggagaggtga cacgttttgt ctcttcatct 540
caacgcatgt tgaccacaga tgacaaagcg gtggtcttga aaaggatcca tgaagtccat 600
gtgaagatgg accgcagcct ggagtaccag cctgtggaat gcgccattgt gatcaacgca 660
gccggagcct ggtctgcgca aatcgcagca ctggctggtg ttggagaggg gccgcctggc 720
accctgcagg gcaccaagct acctgtggag ccgaggaaaa ggtatgtgta tgtgtggcac 780
tgcccccagg gaccaggcct agagactccg cttgttgcag acaccagtgg agcctatttt 840
cgcccgggaag gattaggtag caactaccta ggtggtcgta gccccactga gcaggaagaa 900
ccggacccgg cgaacctgga agtggaccat gatttcttcc aggacaaggt gtggccccat 960
ttggccctga gggtcccagc ttttgagact ctgaaggttc agagcgcctg ggccggctat 1020
tacgactaca acacctttga ccagaatggc gtggtgggcc cccacccgct agttgtcaac 1080
atgtactttg ctactggctt cagtggtcac gggctccagc aggcccctgg cattgggcga 1140
gctgtagcag agatggtact gaagggcagg ttccagacca tcgacctgag cccctttcctc 1200
tttacccgct tttacttggg agagaagatc caggagaaca acatcatctg agcatgtgtg 1260
ctctgcactg gctccactgg cttgcatcct ggctgtgttc acagccttgt ttgctgcttc 1320
catcttcccc agtactgtgc caggccttct ccccctcccc agtgtcctct cctctcaggc 1380
aggccattgc acccatatgg ctgggcaggc acaggcagtg aggccgaggc caatagcgag 1440
tgatgagcgg gatcctagga ctgatctgta gcccatgctg atgtcaccca ccagggcaat 1500
ccatctggag gcctgagcac cctggcccag gactggcttc atcctggcac tgaccaggaa 1560
agactgcctc tgaccctctt agcagacaga gcccaggcat gggagcactc tggggcagcc 1620
tggctcaggt ttattgattt tcgtctgttt accctatcca ttaatcaata catgtaatta 1680
actccttcaa aaaaaaaaaa aaaaaaaatc gcagggggt cccg 1724

<210> 42
<211> 798
<212> DNA
<213> Homo sapiens

<400> 42
tcgacccacg cgtccgactt cggaaactga ataaggtgat tagtgacctg actccagtca 60
gtgagcttcc cttaacagcc cgaccaaggt caaggaagga aaaaaataag ctggcttcca 120
gagcttgtcg gttaaagaag aaagcccagt atgaagctaa taaagtgaaa ttatggggcc 180
tcaacacaga atatgataat ttattgtttg taatcaactc catcaagcaa gagattgtaa 240
accgggtaca gaatccaaga gatgagagag gacccaacat ggggcagaag cttgaaatcc 300
tcattaaaga tactctcgga tttgtcctac agcttagtat tgtggttgac agcgatacta 360
gggctgacag cacagaagtc acaagagaag agtggaaggg caagaattca aagcatttgt 420
tcatacaatg tggcaacctc ttttgcatag ttgcgtagga tcctgtttgt aatgctatca 480
taaatattct gtagtttttt tttttctct cccaactgga gctatgacac tttttattgg 540
attcagtctt gtctcttgtc tagaaagaac tttatcttgt tgacgcatga gctgtttaaa 600
aattatccta ttaaatgttg gttaatagtt gtgcagtttt tcatttcaga tggaaaggca 660
atgcaaattt tgcctttgtt ttctgtcacc ttccaacccc tgagcacttc tagtcagata 720
cagattcatc agtgtatgca acatcctttg taatttaaaa taaaaaaaga tgaaaaaaaa 780
aaaaaaaaaa gggcggcc 798

<210> 43
<211> 693
<212> DNA

<213> Homo sapiens

<400> 43
ggtcgaccca cgcgtccgca aaaaggaatg ctttccctaa tgtcccatct tcatgtccag 60
caacatttgt cttccatcct attgatccta atagtctttg ccttcctaag caatccattc 120
ctaaaccaat aactggcaag aaaaataaga ttaccatgat tggctctagg ggttggcaaa 180
ctttttctta aagggcaaga cagtaaatat tttatgcttt gcagtctatc tggtctctgt 240
tgcaactatt caactctgct gtaatagcaa caatagacaa taaataaata aataaataaa 300
tggatgtggc tgtatttttg taaaacttaa tttacagaag caggtggtaa gcccatgggc 360
catagtttgc taacccctgg aaccttagac cagtaactct gaaacattag catttgttag 420
aatcmactgg ggacttgtta aaatacmaga tcccatcctg gktgctggtt cagcatattt 480
gggatagagc cagtgaattt gcatttctaa taatttccca ggtaatgctg aggctgtgaa 540
ttcaagaacc tcacttaaat atatttctct attgtggggc tatcctgtgc attgtaggat 600
gtttcacagc atccctggcc tctacccact agatgccagt agccaccctg cctctcctct 660
agtatgacga gaaaaaaaaa aaaaaaggcg gcc 693

<210> 44
<211> 1358
<212> DNA
<213> Homo sapiens

<400> 44
ggcagaggga acaaacaaac aaaaaaggca aaattggagc tgtccaggcc aggacacccc 60
gatgtccagg ggcttccata cgaacaggca catgggctgg gaagtacatg aggcccctgg 120
gtagaagggt ctgtcagttc tctcctccct tgccctctgg ragggtcctc ctaacatagc 180
ttccaggagg tgggaggagc agttactgtc agcaggtgtc agccaggtgt cagcttctcc 240
tggggatctc tagatgtctg cttgtgattt ttggcaagta tatgcaaatg agcctcctct 300
cctgccctga gacaagtatc tgcagtgtga acctggcagc ctcagaccca aggggctcag 360
aggaaacttc tctggtttct agagctctgt gctccttcag agaagtcttc cttccttcca 420
gtcagtgtcc ctgtgaagct gggatactca tttcctgtgt accgggcaaa caccggattg 480
ctgattttga gaaatgcctc tcgatggacc tgtaacctgc tggagtctgg gatggtagct 540
gtgggctgga cttggctgat gggatgaccc ggtggctagt gcagcatcac acaagcctgg 600
ttcaagtctt ggctgtgtca tttcctgctg agggaccagg cactgaattt cctacctctt 660
agggtcatta cctatgaggt taaagctacc tcatgggatt gttatacgcc actaatgttg 720
aggcagacac ctcttggcag ggtgactgct catcttagac cctccccttt tctgcgaatt 780
tgggcccctt gatcctctga tgggagctga aaggatgaga ggtgggcatc tagatttagg 840
gaggctgttc aggctttgca ggtcccttac ctgaacacat agaaaccctg gagctgtgac 900
tgtgtccatg tgtgtgtgtt tgtctgtgtg tgttgcgggg gatgggcacc tgcatgaatg 960
tggtagagaa aatggctctg ctcagaggga agatacgcat agcaaggcag ggaccagagg 1020
aatcacaggc gcctggagag cagccgggca ccgcctccag ggacctgccg gcttccctca 1080
gtcctccagg ggcccagcac tcttccttta ggccctgtga gcgtcccttg tcaggataca 1140
ttctctcatt ttgctgaagc tgatttgatt gggtgtctgt ttctcgcagc caaaagagct 1200
ctgaatgagg aaagtgcttc tgtgctaact ccccgcgtct cctgaatttc agtcattcat 1260
gtacccgcct cgaaattttt gcaatatctg tgtaccaact gtccatttac ttaataaaga 1320
agttttcttt aaattaaaaa aaaaaaaaaa aaactcga 1358

<210> 45
<211> 965
<212> DNA
<213> Homo sapiens

<400> 45
caccaccatc tgtacccaca ggggagttag gatcaggcct ccaggtggga gccccasgag 60

```
cagaggaaga ggtggaagag tcctcaccac tgcaagagcc accaagccag gcagcaggca     120
ccacccctgg tccagacccct aaggcctatc agcttctatc agcctgcagt gcctgcctgc   180
tgggcctgtt ggcggccacc aacgcgctga ccaatggcgt gctgcctgcc gtgcagagct     240
tttcctgctt accctacggg cgtctggcct accacctggc tgtggtgctg ggcagtgctg     300
ccaatcccct ggcctgcttc ctggccatgg gtgtgctgtg caggtccttg gcagggctgg     360
gcggcctctc tctgctgggc gtgttctgtg ggggctacct gatggcgctg gcagtcctga     420
gcccctgccc gcccctggtg ggcacctcgg cgggggtggt cctcgtggtg ctgtcgtggg     480
tgctgtgtct tggcgtgttc tcctacgtga aggtggcagc cagctccctg ctgcatggcg     540
ggggccggcc ggcattgctg gcagccggcg tggccatcca ggtgggctct ctgctcggcg     600
ctgttgctat gttccccccg accagcatct atcacgtgtt ccacagcaga aaggactgtg     660
cagaccccctg tgactcctga gcctgggcag gtggggaccc cgctccccaa cacctgtctt   720
tccctcaatg ctgccaccat gcctgagtgc ctgcagccca ggaggcccgc acaccggtac     780
actcgtggac acctacacac tccataggag atcctggctt tccagggtgg gcaagggcaa     840
ggagcaggct tggagccagg gaccagtggg ggctgtaggg taagcccctg agcctgggac     900
ctacatgtgg tttgcgtaat aaaacatttg tatttaaaaa aaaaaaaaaa aaraattact     960
cggtc                                                                 965
```

```
<210> 46
<211> 791
<212> DNA
<213> Homo sapiens

<400> 46
gaattcggca cgagcgttcc tgtgctttcc cttcctgggg atctggctaa aatgcgggtt      60
ctgattctgt aggtctgggg tttccagagt ccgcggtttt gctaagaagc cgcagtgatg     120
ttgacgcggc tggtcctcag tgcacacctg agtagcacga cctctccgcc ctggacgcac     180
gctgccatca gctgggagct ggacaacgtg ctgatgccta gtcccagaat ctggccccag     240
gtgactccaa caggcaggtc tgcctctgtc aggagtgagg gtaacacctc ctcactctgg     300
aatttctcag ctgggcagga tgtgcatgcc atagtaacca gaacctgtga gtctgtgctg     360
agctctgccg tctacaccca cggctgtggc tgtgtgaggt ctgccacaaa cattacctgt     420
cagtcctcag gacaacaaag gcaggcggcc cggcaggaag aggagaactc aatctgcaag     480
gcccatgata gtagagaggg ccgcctgggc taccccctca gtgcccatca gcctggttcc     540
ggtggtccta actagccctg tctccttgcc aatagccctg tgctccccag ccccctcccc     600
catgcagacg gctgctatga catccctgtt ccttaaagtg cggggttcct cgctgccttc     660
tcctccctaa ctggcaccct gtgcaaacct gctgcagaga acagtgtctt gggcagtgcg     720
atagtcctcc agttcaccaa cagtaaaaat ggtctcaatg gggagagaaa aaaaaaaaaa     780
aaaaactcgt a                                                          791-
```

```
<210> 47
<211> 770
<212> DNA
<213> Homo sapiens

<400> 47
gaattcggca cgagtccatt tctggcagtg actttgaaaa gttctgctgc gttacacaaa      60
tacgaaaatc acatattttt ggccttgttc ctctgagaac aaaaacatgt aataagagat     120
acctgctttc atcttttgca atgaatagaa tactctccta cttagaaaca ggcttttttct   180
ccttgccact ttattttttc cttacctatg aattgcatgt gcctttgatg aaaacaatga     240
actggacatg tacaacggta catgtaatag actgaatgca acttagaagt ggccactctt     300
ccagtgtaca taggcttgga aatgaactaa tccaaacctg agtaatttgt ttatagtacc     360
tcctttcact tttgtttatt ggtatctaca gtctctcatt ctttttcttt aataatatct     420
ctttatatag aatttttatat tcagccatga ctctattatt tcaatagtca cattaccact  480
tcgaggattg ataccatgaa aaaaggttat ctagtagttt tgagtgaaga tacgaggcac     540
```

accttcaata ccaataagaa ggtatacaac aaaggtctaa tgaagaaaaa tatctcattt 600
tgaaggtagc acatagcttt caactgactg ggcctgttat ggtctttgct gtgtttgtta 660
tcacagtatc taatagtgaa gtggtaatta ctttctttag tagaaattcc aagatctaaa 720
ttggtacaca tataaatatt tgacaacaaa aaaaaaaaaa aaaaactcga 779

<210> 48
<211> 875
<212> DNA
<213> Homo sapiens

<400> 48
gaattcggca cgagctgggt cttctagaag acgaagatct atccaaaatc aagaagcctt 60
tgatttagat gttgctgtaa aagaaaataa agatgatctc aatcatgtgg atttgaatgt 120
gtgtacaagc ttttcgggcc cgggtaggag tggcatggct cttatggaag ttaacctatt 180
aagtggcttt atggtgcctt cagaagcaat ttctctgagc gagacagtga agaaagtgga 240
atatgatcat ggaaaactca acctctattt agattctgta aatgaaaccc agttttgtgt 300
taatattcct gctgtgagaa actttaaagt ttcaaatacc caagatgctt cagtgtccat 360
agtggattac tatgagccaa ggagacaggc ggtgagaagt tacaactctg aagtgaagct 420
gtcctcctgt gacctttgca gtgatgtcca gggctgccgt ccttgtgagg atggagcttc 480
aggctcccat catcamtctt cagtcatttt tattttctgt ttcaagcttc tgtactttat 540
ggaactttgg ctgtgattta tttttaaagg actctgtgta acactaacat ttccagtagt 600
cacatgtgat tgttttgttt tcgtagaaga atactgcttc tattttgaaa aaagagtttt 660
ttttctttct atggggttgc agggatggtg tacaacaggt cctagcatgt atagctgcat 720
agatttcttc acctgatctt tgtgtggaag atcagaatga atgcagttgt gtgtctatat 780
tttcccctct caaaatcttt tagaattttt ttggaggtgt ttgttttctc cagaataaag 840
gtattacttt agaaaaaaaa aaaaaaaaaa tcgaa 875

<210> 49
<211> 614
<212> DNA
<213> Homo sapiens

<400> 49
ggtcgaccca cgcgtccgac ctccccctcc tgggctaaag tggttctcag ctcactgcaa 60
cctccccatc ctggctcaag tggtctcgtg cctcagcctc ccgagtagct gggacaacag 120
gagagcgcca ccaggcctgg ctaattttgc atgttttgta gaggcagggt ttcaccatgt 180
tggccaggct ggtctcagac tcttgataaa ataaatgatt aattgtggca ttttggtttt 240
caaaatgaga attgtgttta aaatgcaaaa gagggaaaga aagttatatg taatcttcct 300
atatttagct tttattttac ttcattggca gtctgggtaa aaaattcata gaagacagaa 360
gacttggttt ctagtcttgg cctgaaactt ttagctgtca caactggggg atgctgttgg 420
catctagtgg gtggaggcca gggatgctgc aaaacattcc acagtacaca ggacagcgcc 480
ttacaggttg aaggtttata caaataatat taaagctctt tttttatatt aatgtggaaa 540
aatgttattt tggttcccat gagaaactgc tactatttga aatttaaaaa aaaaaaaaaa 600
aaaaaagggc ggcc 614

<210> 50
<211> 556
<212> DNA
<213> Homo sapiens

<220>
<221> SITE

<222> (513)
<223> n equals a,t,g, or c

<400> 50
gaattcggca cgagaagatg ggcagccaat ggtgctcaaa ctcaaggact ggcctcctgg 60
ggaagatttt cgagacatga tgccaaccag gtttgaagat ctgatggaga accttcctct 120
gccagaatat accaaacgag atggcaggct caatctggcc tctaggctac ctagctactt 180
tgtaaggcct gatctgggcc ccaagatgta caacgcctat ggtatgaggg agaggctaaa 240
attgctcttt tgggggactg ttgttcttat ttcaactata gaaggatatc tgtggtcaat 300
gtcaggtata gagatgattg caggcaagtg ctggagaagt gaatagtatc caaggtggtc 360
ttgaatatgt ttgcttttgt catattggtt ttcataacat ccatgtgggc ccagaccata 420
agcttacatg tctccagtag tgaggaagtt tcctgttaag aactctaccc aaggagccat 480
attctcgaag gggggggccg gtacccaatt cgncctatag tggagtcgta ttacaattca 540
ctgggccgtc cgttta 556

<210> 51
<211> 1003
<212> DNA
<213> Homo sapiens

<400> 51
gaattcggca cgaggtcggt gagcatcggg tactggatgg ggatgttgcg atgcgcctgg 60
gcgttggcgc cgcctgttcc cccaccactg gtcactgact tgccgttctt tttcaccttg 120
tcgccgttct tgttcgcctt ggaaccgcct ttgccggact tgaccgactc agcatcgatg 180
tccgtcatcg tagatcgtag atctcgaggc tctgatacca attgttggct tttaaaccgt 240
agatcaaaac acccaggagc accacgtatg tgtacgtgca aagctaactc gaacaagtac 300
actagcagct tgaccgatta gcccttgtac acacgtatgt gcaactagct agagacttgc 360
gtatgaatac ggttcagccg actagcttcg gttgattgga tcaatcacgc ggcaatggat 420
caactcggst ctctcaacaa gaacgtaaaa mgcaargcac tgaatcgttg atggcacagg 480
ccgtgtcgag cttctcgtaa tactttgtat tgctttgccc atctgatctg gtgttacaag 540
caccacccctt agggtgcctt ttatacacat cccacaaggg actatgggcg gtgatgaaga 600
tgaagattat tctccgaacc ctcctagtgt ggcacgcaat cacggacgac gacgtcgatg 660
acgactccga cgaaggtgcc atggccgcca tagcccggta catgccggat tccgtgctga 720
tgacattggc ggagttcgag acagcaagag aggcgtggaa cgcactcaag aagatgagga 780
tcggagaaga tcgcgtcacc aaggcttgga cacaagtgct gaaacgccaa tttcacaagt 840
tgcacatgga ggaaactgaa tcggtgaacg actacgccat gtgtcttact actttggtgg 900
gagagttccg cgcgcttggt gcaaagctcg atgagaccga gattgtggag aaaattttca 960
gttcagtgac tgacaaattc acgtacatca tcggcacgct cga 1003

<210> 52
<211> 886
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (92)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (886)
<223> n equals a,t,g, or c

<400> 52
atttcatttt agggcatact gggcttactc tcctcccagc tgtctgtgga ttgatttgat 60
tttaatgttc gagttttaca gcaacagctg anaaaccatg aactattcta ggaactgtgt 120
tggaactctt taaaataaag aaaagaggag gaggagagga agaaagaaaa ccaacttaag 180
aagccttgga ctttggaggg acagaaagcc accagccaat ggagaacaaa gagatgtttc 240
cctttccttt ctttcacctt gtcattctgg gtttccttct gcttcactct ttccttcccc 300
cttaaaagtg gtattcctgg ttggtctgtc tgtctgtcct tgtccttgtg gtgatcctgg 360
catggtgata tgctccactt tgcattatcc atggtctctt accagcgcac aagtcagtgg 420
ggaggatcta accacgcctg gtggtgagga agctgaattt ccaggcctgc gtcccatgta 480
gcctctccat gaactgcaga aggcatgttc tgcatggtta ccagtaagtg gctccctctc 540
accgtgttca ttgtcaaatg agagcaaact ttaggtgttg gctccattgt acactctact 600
tgctctgctc ccctccctcc aaccagggtt catgtcagtg cacaccccat gtgccctggc 660
gaagctggtg ctgtgagtga tgtttcccat acaactcagg gatgccaggt ggcttaccct 720
gagatagtca ttttgggcac ataacagtgt aggaatgaaa catggatttc attgatattt 780
aaatctgtca atttcatttt ttgttaatgt tttcccctga tgacttttta gcaatttaac 840
aaataaaatg gacaattgtc ttaaaaaaaa aaaaaaaaaa ctcgan 886

<210> 53
<211> 564
<212> DNA
<213> Homo sapiens

<400> 53
tcgagttttt tttttttttt tttttttgag acagagtctt gctctgtcgt ccaggctgga 60
gtgcagtggc gcgatgtcgg ctcactgcaa cctccacctc ccgggttcaa gcaattctcc 120
cacctcagcc tcccaagttg ctgggattac cagagaagag gctgaagggc aaggagggaa 180
aggaattggt tcccaggtcc atggacctct tgtgaagccc ccattgctgt ggggtctgag 240
gaaacacaga ggaggtgtca gctgctctgc ctgcccccac tcccctgcca acaacgtagt 300
aacctctgtg cctaacctct gagccctggc ctccaaccct gggagggagg tacttatgtt 360
atccgcattg tgcacgtgga gctcagaggg gcagccactt gccaggccag caatccaggc 420
tgtctgtctc cagagcccag gcccccagtc aacaacttgc caggtgcccc tctccaggcc 480
tcggcttctc cacctgtggg tcaagagcac caggcttgtt ctagagctat cttctcagac 540
ctgatgtggg ctcgtgccga attc 564

<210> 54
<211> 933
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (425)
<223> n equals a,t,g, or c

<400> 54
gttgggtttt aatcctgggt ttgacactta ataaactgca tgatcttgga aaagacatgt 60
aacctctctt tcaatttcct tatgtgtaag atgcttataa tagtattcac tttgtagata 120
ctattattgc aaggactaaa attatgaata tgtgctggca aataccaaac tttatattaa 180
tacaagtgtc atcagaatat gtacatatat taatagtaat tgttaccaaa acaccagggg 240
ttcaatctgg gtcctgctgc tcactgcaca gaaagccaat gcctgagaca acaagtgttg 300
ccaaggaaga aggcttaatt gggtgctgca gccgaggaga tgggagctca gtctcaaatc 360
catctctctg acagaccaaa actggctata tagcarggaa gaaatgtaat catgtgtggg 420

aaaancrgga actcagaagg ggcttggaag caatcatgtt gaatcagcgt ccacatttta 480
ttgtctggat gtgatctggt gagtttcatt tctttgatac tttttttgag aggcctgaag 540
gtcatttcct gaggaaggat ctcagataaa acaaatataa gtttcaaatg ttaagaccag 600
aaagttcaat ttctatgttt atttattctt tttttaaaa aaaaaagcta tatgggactg 660
ttgggttggt ttcataatgg ctgagtactt tgaaggttct gtggttgcat gaatggagaa 720
gatagagtga tgggtgggga ctttaaaata ggatgatcca ggaatgccct gaagtagaga 780
cttgtaagaa tgagaaatag caagttatgc gggtggcata gaaaaagctt ccagattgaa 840
aagcaaaggc aaagaggatg tcttgtgcag agaaagcatt tgacaaaatc gaatgcctct 900
tcatgttttt aaaaaaaaaa aaaaaaaact cga 933

<210> 55
<211> 597
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (12)
<223> n equals a,t,g, or c

<400> 55
cttgccatat ancaagctga attacctcat aaggaacaaa gtggagytca cgcgktgcgc 60
cgtctagact atgatccccg gctgcagaat tcggcacgag cagtccagaa actgcgtgcc 120
ctgccctttg cttgggcccc tctaccagta tgtccagcat gtgcccgggg gccctcagct 180
cccctggggc ccagcccacc caagacacag ctcttggtcg tgaacatgaa gatgagccaa 240
actctagtgg ctcttcctga aagaaatgag aatgcccagc cacacccatg cacgctttgt 300
tctttttat ttaatactga ggaaccggag tggaggggtc ctgccgggct gcagtgaccc 360
tgagggaagt caggagagcc ctgggctgca gaagagtccc cccacaggct ccgaagcaag 420
cttgtcctgg tgcattcaga ctgctcacag caggctttgg gccctcactc tccagatccc 480
agagagccct ccagggctcc cagctctcgg gccagtgccc amgtcctcga aggggggccg 540
gtaaccaatt cgccctatag tgagtcgtat tacaattcac tggccgtcgt tttacaa 597

<210> 56
<211> 773
<212> DNA
<213> Homo sapiens

<400> 56
gaattcggca cgaggaccag gcccctgcga tgctcccaaa gcctcagctg tccgtcctca 60
cactcactgt ggcgctcagc ytcatcccag gaacctgact gcctgtctcc ccaggcgaag 120
gcttcatgag caaagccact gcagcatcgc acggtgtatc tctgagcaca gctgacttga 180
cagaaggact caactgtcca cattaccgar gactgaggta tacggaatgg tttctgtttt 240
gcttcttcaa ggaggggaac tgaaacccaa ctaaatccaa ggtgcctctt ccaacgcctg 300
taactaaact tcaagcatca cagccccaac acctgctgat ggcaccattt taactgaggt 360
ccatcccgca agcttcccga ctgtccacac tggctctctc tactcctgtg caccaaagar 420
acaagccaga ataaatggat aaaagacagt gtatgcgcat gcctgtccca gctacccagg 480
aggctgaggc atgagaaccg cttgaacccg ggaggcagag gttgcagtga gccgagacgg 540
cgccactgca ctccagcctg ggagacagag cgagactcta aaaaataaat aaataaatta 600
aataaataaa taaataaaat taaaaagata gtgtaggcta caaacctcag gaagaaaata 660
ccagcatgac ttcagaatag tcagammtaa tggtgtataa agttctcccg gctcctctcc 720
acccacctcc atcaatccca ccctatctct aacccccaag ttctctgttc ctc 773

<210> 57
<211> 733
<212> DNA
<213> Homo sapiens

<400> 57
gaattcggca cgagctggct tgcagcagaa ctggtttttt ggccctgagt caatgttcat 60
ttccccacac caccttaaca gggtttcctg gccaaagagc agggatggag atgatgatgg 120
tggtgatggg gtgtgtgcaa ggcccgggag agggttgtag tgggaagatg gggaagaagc 180
cacgcccctg gccactagtt tcttattcga ttactcatct gtagagaaat ttgagacgca 240
tcacctgacc catccgtcaa ttcgcatctg gcatctaaaa gcaccagagt cagtgctggg 300
gaaaacacta tttaaaaaaa ttcccagttt aacctcatta agcctctgtt ttcccatttg 360
taaactacag acagactgga gacttgtaag agataaatct aattctttca tagacattaa 420
tgatccttga aaaaggatca tttgagggac atggagattg gtttctactg tttctgttgt 480
tactaacact cctcctttcc caaggccttt agaaaggggt gagctctcca tcacagaaag 540
tattcagata ggcttccagg aatttttggg gaaaatgttc ctgctttgag taagacacag 600
gactagatca gcgtttggca aactatggct cgtgggctaa attccgcccc tctcctgtgt 660
ttgaagataa agtggaacac agccacgttt actcgttgac agagtctacg gttgctttgg 720
cacacagact cga 733

<210> 58
<211> 531
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (506)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (520)
<223> n equals a,t,g, or c

<400> 58
tggcggcccg ctctaggaac tagtggaatc ccccgggctg gcaggaattc ggcacgagac 60
ttccacaaac tcttcattgt ctactgacaa ccttacttct atctttactg agaccaaaaa 120
aaaaaatcag atgagttatg cccatcacgt caccgtattc ccaaactacc tgcctctgtg 180
cacaccacct cactgcctgc tgcagttact gtccagggcc agcgcctctg cccatgtact 240
ggagcctgtc cctccacccct tttcaagcat gttactctat caaataaata tccctttctc 300
ttttgcatta tcagttttgc tatctctctg ttggccccac cagcactatt acccatgcta 360
tattagcttt taaaaaattc tctcaatctc acatttatct ccaacgttta catcattctt 420
ttgctgcact ttgtagaaaa atattttgaa ttttctgtat ctatttctac ttccttactt 480
cccatgtttt cttgaactca ctcganggggg gggccgggan ccaattcggc c 531

<210> 59
<211> 852
<212> DNA
<213> Homo sapiens

<400> 59
gaattcggca cgagtgaact gcatgtccat ttatcttaag ccaacacctc tacttatgta 60

```
ctagatccca ttctttcttc tccccttct ctcttgtatt agcaacattt ccttctttta    120
ctgtaccgta taaagatgct atatttctc ccatctttaa aaaagaaaaa gtctctttta    180
accctatatc tccctccagc tactaactgt atwkctctct tgtgctttaa agaaaaaaaa   240
atgtgtgtgt gtgtgttttc tttttgtttg ttttgtttgt ttgtttttgg tatggtctta   300
acggtcttgc tttgtcaccc agggtggagt gcagttgtgt gattgtggct cactgcagcc   360
tcagtctcct gggttcaccg ggctcaagtg atcctctcac ctcagctcct gaataccttg   420
gaatacaggc atgtgctgcc atgcctggct agagaaacgt tcttgaaacg tttcatatac   480
ttaatatttt taattccttg ccttccattc tttcttgaat ccactccaat cagattttta   540
ttcttgccat tcttctaaaa ctactcctat gaaggttatc tgtggccttc atttttgtat   600
gtttactcca agaaaattgt tgtgataaat taccccagaa tgtagaggtg taaaacaact   660
atttattacg ctcatgagtt ttgtgcatta gaaattcaga caagacacag caggagcagc   720
ttctctgttc cacagtatct ggagccttgg cttgaagatc aaagcctagg ggcttaattg   780
tcagaaatga tcgtgtgtat gtctgggagt tgataccagc gtttgtctgg gaacctcagt   840
tcctttcctg cg                                                       852
```

<210> 60
<211> 680
<212> DNA
<213> Homo sapiens

<400> 60

```
gaattcggca cgagaaaaaa acaaaaatat gttaatattc tgtggagaat attggtattt   60
ttgttttaat cttctgtggg ttgtggttcc atataaattc agttttctga gctttggcag   120
tgttattcag atctgtccca caagtgttcc acccattggt cagtctggga tctgggtgta   180
ggtctactca ttatctcagt tatcagagtt tttattatgc caattggtat cagatgcata   240
cctacacagg ttgaggatga gcccagcagt tcataaacaa cattatgggg tcactttcct   300
atggacagag agagaagaaa aaaaacccaa aacaacagag tttgtcctgc ccacttggag   360
gcacggctcc acaggatgga gagaaaggtt cccttccttc aaaagttttg ttcctggagg   420
ctttccattc ccagattctt ttgttgttgc tgctgccccc accatggatg acctggggac   480
tgacacatga gagtatggag ttttcccaag ctgctgagca cagtggctca cacctgtaat   540
cctagcactt tgtggggatg aggcgggarg ataacttgag cccaggagtt tgaggttgtg   600
gmgagctgtg attgtgccac tgcattctag cctgggcaac agagtgagat cctgtcaaaa   660
aaaaaaaaaa aaaaactcga                                               680
```

<210> 61
<211> 894
<212> DNA
<213> Homo sapiens

<400> 61

```
tcgaggttag actgcataga aaacaatttc agatttcctg gaggctgcat aaaatttaac   60
tatttaaaga taattaaaga agcattaaaa ataagaagat tatcatctcc agcaaaatat   120
agaaagtagt acagtgaaca aaatataatt agagaatttt tgctcaaaga aaaccatctt   180
tacattgtaa caggaaaaaa tgtgtgtggg ttttaccaaa tttttattta gaaatgataa   240
ggaaataaga agtctaaatg gttccaaatt ctagtatgtc aaaataggaa atcaagtgat   300
aatatctaaa agtgatgaat caacaaatag ctttagtcaa tggtatttac atacatagaa   360
ctaaatatta gaaggaacaa ccaaagaatt gaacatcttt gcctgtgaag agtcacttag   420
ggattcgaag ggaaaagcag actgatgctt tttttgtctc agcactatgc gatttttaaa   480
attgttttcc cacaatatat tgatacaact aaaaattatt ttaaaattaa aagtttcttc   540
agtgctcccc tctgtcaaat ctttaaaaga tgaaagaatc atatttattt tccaagtcag   600
tctaaacaaa gttttaagtc catgcctgag attttatcca cagcgtacag caacatttct   660
gtcttgccaa attgagtttg ttcagcagct tagaaacact ggcaagatac aaaactagtg   720
caagcatatt ttatttaaaa aatagtcaga caacatcttt caaacaccat tggttagttt   780
```

tcatacaaaa tgcaagtttt atcagggtat atttttattg taaacttttc aaaattattt 840
ttaattatgt gggcattttt tatgtctaac tttatttgca ctcgtgccga attc 894

<210> 62
<211> 691
<212> DNA
<213> Homo sapiens

<400> 62
gaattcggca cgagatccta ctatatttta tgaataaaga ataaaaatga gtgaagcctg 60
atcctccaag agcaccagga gaaaatgaag attctagtgt tccagaaact ccagataatg 120
aaagaaaagc aagtatatca tatttcaaaa atcaaagagg aatacagtat attgatttgt 180
cttctgatag tgaagatgtc gtttccccaa attgctccaa tacagttcaa gagaaaacat 240
tcaacaaaga tacagtgatt atagtttctg agccatctga agatgaagag tcccaaggcc 300
ttcctaccat ggcacgtaga aatgatgata tttcagaact ggaagacctt tcggaattgg 360
aagaccttaa agatgctaaa cttcagactt tgaaggaact ttttccacaa agaagtgaca 420
atgatttact taaggttata ttcattggtt attgtagctg taatgatgat aaaatctctc 480
ctgcattcag tgctatagtt agtagtggat agtcattttt ctaaagatat cttacgtttg 540
aagatattaa ctattaaatc taaaggaagt aaatgccaga catttattta ttgaaagtct 600
taacttttta atagatgagg ttatttattt gtaaatagtg cagtaattaa agccttaata 660
gcgaaaaaaa aaaaaaaaaa aaaaaactcg a 691

<210> 63
<211> 891
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (14)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (398)
<223> n equals a,t,g, or c

<400> 63
agtgaaataa cacnttacct gaaggtcagg ttctgaagtt agcatatgag atgaaaattg 60
cttatggtta aaatatcctt tcagagcctt ggaagtcacc agtaagcagg gccagatgca 120
cctggtttgt ggggcatgtg ggatcaaaga cccactaaag gaacacagga ttttcagctc 180
cttttgctcc ctggcatttg ctcatcattt gcactattac taaatgctct tcccttccct 240
gctccttctc caagcattgg tacatgtctt tgtgctagtt aagcttgagt acattgtgat 300
ttcactagat cacactccca atttcaagkk cagtgtgaag aatatagagg ttctggttgg 360
tctagccttg gccacgtatg agtagacacc cccagttnca aaggtcaact ccacttctca 420
ctagaattaa aaagctttac tccaaatgta gttaaaacag cccaatatct tcctcttata 480
agcagtaatt aaactttagt gtggataaga ttcatctggt ttgcttactt gaaaatgcag 540
atctttggct caacctctag aagatgggac agagccagag tggggttgga tggggttgag 600
aaatctgcat ttcaacagta gtccacaggt gactctatgc agaccctgga aaacactcta 660
tttaagggct caccacagcc agggaccata ttccaactgt cacttttcta ggtctcattc 720
tcattatttg ttccaagact ctctcttatt tttgcaaatt taatttaaaa gtatgagcat 780
ttcctgaatg taaccagcca ctctaagcca gagctgacct atgagggaca catacgtggc 840
caaggctaga ccaaccagaa cctctatatt cttcacactg aaccggcacg a 891

<210> 64
<211> 958
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (469)
<223> n equals a,t,g, or c

<400> 64
gaattcggca cgagcgccca cctagtgcac agccttagag gtgttacagg taaaggaatg 60
gcggctcaga gggaaggaga gacttgccta acttcagggc aagctaacgc ttgatttcaa 120
cttgataaat ttctgagtat gcagtgggtg cacatagcag agacaggtaa tgagaagttt 180
tcttttttcc ttttcttttt ttgtgggggg tggggacaga gtctcactct gtcacccagg 240
caggagtgta gtggtgcaat ctcggctcac tgcaatctcc ccccaccccca cctccaggtt 300
caagcgattc ttgtgcctcc gcctcctgag cagctggcac tacaggtgca cgccamcacg 360
ctgggctaag ttttgtattt tagtagggat ggggtttcac catgttctac gtttcaccat 420
gttggccagg ctggtcttgg actcctggcc tgaagtgatc tgcctgccnt cagtgtccca 480
aaagtgttgg gattacaggc gtgagccacc gcactcggcc gagaagtttt tctgattaaa 540
aaaaatttta aggcacacac ttcagacagt ggctgtgaag gaaccctgat gtgtatctaa 600
actgtcgcct cgtgcacatc accccattac ttactctgtg ctaagtgctg tcatgcatta 660
catcattact ccttagaaca ggcctatgag gtggagtctg cattaggccc attttggaca 720
aggacaccaa tagtgtggga ggtggtgtac cttgcccaag cccccagcag gtaagtggtg 780
gtggggatta ggacccaggt cacttgagtc catatcctgg gctcttagtc ccactctgcc 840
tggctgcctg ctgctccatg aagccaaccc tggacctaga cctggacctg gatcgtcata 900
gcccagatcc ctgtgtgctt cccaggctgc cttgtggcag gtggatggtg cccctcga 958

<210> 65
<211> 802
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (291)
<223> n equals a,t,g, or c

<400> 65
gaattcggca cgaggaaata tctgaaaact tacatctgtt cttgtgagac acatcattgc 60
tttgacgtca gtgacctaac cacacgaact aagaattttt aaaaagtact cttgatggta 120
tgttttatga tgttcagggt cccagccatt cctcggaaat gtgttttttt gttttttttt 180
ttgtttgttt gtttttgttt ttgatgaatg agtctaaagg ctgagtggct atcaaacaat 240
tctttttggt ttacattgta ttatgaaaat aatataaaaa ccctgtgtac ntttcttgtt 300
ttcctttcta tagttttggg gaacaggtgg gtttttgkta cctggataag tctttagtgg 360
taatttctga gattttggtg tgcccatcac cccyccgtgt actttaaaat gagtaagttg 420
tgaaaatgtc aactagtttg ctatttagag ggtcctcata aagtaacaaa atgatacata 480
acacatttgc acagcaagtc ctcacttaga gttgtagata tgttcttgaa aactgcgact 540
tcaagtgaaa caacatataa caaaactaat tttaccatag gctggttgac acaaacaaga 600
gcttagttcc taccacacat tactggtcat aaaaacatga ccaaatctct aactaaagac 660
caaaagactt ctaataataa acatcgagat aaatgtgagc tatacctacc tttaagaaag 720
attagtgtaa acaagtaagg taatttactc agttattcta gttcaggact gtgggtagcc 780

agagcctgtc ctggcagctc ga 802

<210> 66
<211> 1092
<212> DNA
<213> Homo sapiens

<400> 66
gtcacacggt cgaatagctc cttcttctca gtaatacaag cttttttggt ttgaaatatg 60
gatcctcttc ccagcataat aatgtgattt tttttattca ttttatgtta ttatattcac 120
atttttactt aaaggaaaat gctgctattt gtgatgaaat tgctcgtctt gaggaaaaat 180
ttcttaaagc aaaagaagaa agaaggtgag ctggcttcat tttgtgttca gcatcacctt 240
tttggtgatt gatttggtga ttgataatgg tgttactgct ctggagactt tttttcccag 300
tgggattgat gcgtatcgca cagccccttg gccacttgat caagcacaga gaaacttaca 360
gcctgaggca ttggtgcctg cacacccaag ttatgttggg ccatggcgat gagacagctc 420
ctctactcat ctttctgaaa aagccatctt gccacatcta ataaataatc ttactaagat 480
tatttaatct tatggcccaa ttataaaagc caagtgataa aagcaactgc ctctcgttct 540
acaaatattt attctgtacg tactattctg tgcaaagcac aatgggtata tatacatgtg 600
taaataatgt gcctttcaga agcctaacac cgtccaacat caaggtagag gaaccgtcca 660
gatgcaagag ataagctaca gttcttatcc ttggcctctt gaagtattga ttatcctcca 720
gggctttatg attcataggg cctaataaga acctttcttt tatgagtata gtaatctttg 780
tatataattc tggcttttcc cagtacttga gtaaaatact gaattgagac aatacggaag 840
ttcatttctc tgctcctttc cttcctgatc tcaggtactt gctaaagaag ctcctccagc 900
ttcaggctct aactgaaggg gaagtacagg ctgcagctcc ttcccacagt tccagtttgc 960
ccctgactta tggtgtggcc agctctgtgg gaactataca gggagctggg cctatttcag 1020
ggcccagcac tggggctgag gaaccatttg ggaagaaaac taagaaggag aaaaaaaaaa 1080
aaaaaactcg aa 1092

<210> 67
<211> 734
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (396)
<223> n equals a,t,g, or c

<400> 67
gaattcggca cgaggtgaat ttaatttccc ctaatgactt atggtattga gcatcttgtc 60
atgtgctcat tggccattta tagatctact ttagagaaat gtctattcaa gtcctttgcc 120
cattgttttg ttttgcttca ttttttattt taggttcaag ggtgaatgt gcaggttttt 180
acacgcatgt attgcaagat cctagagctt gggcttctaa tgatcctgcc acccaagtag 240
tgaacatagt acccaatagg gagttttcaa cgcttgccct ccttctccct ccccactttt 300
ggaatccctg gtgtccactg ttcccgtgtt gtgccatgtg tccccagtgt tgagctccca 360
cttatgagtg agaacatgtg gtttttggtt tctgtntctg cattaattca cttaggataa 420
tggcccccag ctgcatctat gttgccacat tgtacatgat ttcattcctt tttctggctg 480
tgtagtattc cataatgtat atgtaccaat tttcttttct tgtcttttca gagacagggt 540
ctcactctgt cacttaggct gaagtgcagt gacatgatca cagctcattg cagcctcaac 600
ttcccaggct caagcaatcc ccctatctca gcctcctgag tagctgggac tgcaggtgca 660
taccaccaca cctggctaat tttgtatttt ttggtagaga cgaggtttca tcatgttgcc 720
caggctggtc tcga 734

<210> 68
<211> 706
<212> DNA
<213> Homo sapiens

<400> 68
gtttttgtgt atctgtctta ggcttttta tttgaggtta ccattaagct tgcaaataac 60
atgttataag ccattatgtt aaagtgatga cagcactgat tgaaaaagaa aaaaacaaat 120
taacaaacaa gcacagagat aactaataac actacattta attttattcc cctttttaac 180
tttttattta tttatatatt atagtgctat gtcttgaaaa gttgttgtag ttattatttt 240
gataggttta tcttttagtc tttctacaca agatatgagt agtttacaca ctacaattgc 300
agtgtcataa tattctgtgt ttgtctgtga gtwttgtacc ttcagacaat ttcttattgc 360
tccctttctt ttcagaatga agaactccct ttagcatttc ttatagcata ggtctggtgt 420
taatgaggtc cctcagcttt ttgtttacct gggaaaatct ttatttctct ttcacgtttg 480
aagtctattt ttactggatg tactattcta ggatgaaagt tttttccttc aacactttaa 540
atatgttatg tcactttctc ctggcatgta aggtttccct gagaagcctg ctgcaagatg 600
tgtgggagct catttgtatg ttatttgttt cttttctcty actgccttct tttaagattc 660
tttctttatc cttgaccttt gggagtttga ttattaaatg cctcga 706

<210> 69
<211> 436
<212> DNA
<213> Homo sapiens

<400> 69
tcgaggggcc cggccctcgg cgtcccccag gctctcaccc gaagccgccg ggctccytcc 60
gaggtccccg cggtytccgg tcccctcttc cggaggcggc tccaggtgtg cggccaacac 120
aggtgaaagg gscggggccg cgggaggggc cggggcgctc cctggctgcc tgaatggccg 180
ggcggggtcg agggagagtc gcttcctcct gggtgggggg cactggccca acctgctgtg 240
gttgcaaatg gcccggccag ttaactgagc atctactgtt tgcagatcct acattgaggt 300
agcctccgct cctttcccgt cacgactgcc ttgccctgtg gggcaggaaa ttattagcaa 360
tgacaacaac accgaatctg acatcttaag cattctgcta agtaaactct tttttatttt 420
ttctcgtgcc gaattc 436

<210> 70
<211> 721
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (7)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (644)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (718)

<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (719)
<223> n equals a,t,g, or c

<400> 70
ttatcanaga ccctgtgggg gaatggggtg gactctgggg ggttagcctt cttccccagg 60
ctgggagtgg gtgagacgag actcggggcc tctacatctg agtgtccccc aaaccgagca 120
gtcatgtcgc gagcaaacaa agaaatcatg ttacttcttc cagctgatgt tccacttgtt 180
tattctgttg tttctgtggg gagagtcaca ttaaggtgat ggagggtggc cccctcaact 240
ctattcccca gagcaggaag tggtaggcag gggccaggaa tggattttaa aggcaaagtt 300
ctcagaccca gtgggaactc gaactggtaa actctcctca agctcccaag gacagaggat 360
ttgggtcttt gttggctttt gtccacagcc acagaactca aggtctgaat ctggaatctc 420
ttgacaggac agtaacataa acctctagag atggagtttg agaaaggccc ccccttctgc 480
cagcttgtga tttagaaaag tgcattcatt caataaacat ttactgagca cgtacgggcc 540
aagtacggtt cttcacagaa gatttagggc ggaaaaggac agacaggagc ctttggccct 600
gaggtttcca ttctaggagg cctttaaatc tcagactctc agantaacag agactatgat 660
tactcactat tcctctggaa cacgagccaa aagagagtgc tgtcagatca agacaatnng 720
g 721

<210> 71
<211> 793
<212> DNA
<213> Homo sapiens

<400> 71
gattcggcac gaggtttcat gttaattttt tgttttggta ttgcctgaat catacaagta 60
gtgtaagttt ggacctcaca tcaagagaaa taacagaatt ggggacttaa atttttaatt 120
ttaaattttt ttctacctca catcagatag agacaagcct cattgccatc tccctgtacc 180
agaatgtgga atttttcttg ttcaaccagt atttgtgagt atggcttttt aaaatttctg 240
gttttatatt tactttccac ttctatgtct tcacctctta taggcccaga acctcactct 300
ccaaccaagt gcaaaattaa atagaattct tgtgatatca ggggaaacaa aatatctccc 360
tgaccttctc tagatgtctg tactatcagt tcacgagttt ctgtctctaa agcatagtcc 420
ctgtttctcc tgatgttttc tctctttctg gcaaaaaagr atgttattgc atattacaaa 480
taatttttgt tagtttctac tcaaaatttt aacatatttg tagtgagaaa gatgttacaa 540
tatttatttc accatcctgc cagaacaaga tgtcaaggtg gtgttttctg aaacacaaat 600
gggtgtgtca cactcctact taaaatcttc aatgacttta tatttctatt atcataaaat 660
tccatctcct tcatattaca taaaaggaaa tcctaccttt caagtctaac cctttgctat 720
ggtctccttc acactcagtt ttcagcwata tggagctcct ttcaggtcct aagatgtttt 780
ggtgtttcct cga 793

<210> 72
<211> 761
<212> DNA
<213> Homo sapiens

<400> 72
gaattcggca cgaggacagc atgagatggt ctttgttgtt cttctcccgg aaatgattcc 60
cctcactgca gaggagggtg gaggctggaa gaaaagcaga agtgacccta agacacttcc 120
tgtccaggca tttgttttca agtgccaagc ctggggaccc aggaggagaa gggaaggact 180
tccctgggat tcctccaaac tgtctcccct gagcagcact agactcacta cctgctcccc 240

acctcccacc tcaggaaggg gactgcaggg tacacaggag gctgcgccct ggacaccagg 300
ccccagcccc accaaaccct cagtccccaa agccccagac cctgaacttg ccaggaccat 360
gcaggctggg ctactgtggg tcttggcaga accagcaacc aatggagggc gagaaggaag 420
gagatctcta acattttcac agaacaaacc acgcaggaac ccaagaaagg ctgaagttct 480
attttttgcc aatccggtgt aatgagagta taaagccaaa attaacttga attctagaaa 540
ataaagacaa gccatatttc ctgaacctga gtcaatggac tgagattcca tccaaataaa 600
ggaaaggcta ggagggagac gggtggcttc tggctccagt gagacccgag gctatctgct 660
gcagacccca gattgcaggc cacggtccct gtccagtggc agggcaccag cctaccttgc 720
cactgtgggc agccatcagg gagagggcag ccactmtcga c 761

<210> 73
<211> 673
<212> DNA
<213> Homo sapiens

<400> 73
gaattcggca cgaggtccta gtgtgtgaat caggccctgt gtggacatgg tcgtgccagc 60
ggactcggga ggcctgccgc gccgcaccga gaagctgctg tgtgtgatgc ttttgcttct 120
ggagaggatg gcactgtgcc ctgtgcttga tgtacacaca catttggggt gcatcatctg 180
tgtggcctgc cagcctgtcc gcactgttct gtctcttctg acagcctcca tccaggaagg 240
ctctagacta tctgggcatt ttcaaacact gccgcatcaa actgatacaa ctttccacaa 300
aggaagcaaa ttatagagct gagaccaaac cagttttatc ctcctccctt accccacccc 360
cggcatattt tgaatcaaac aaactcttct tgtaatgtcc gctttccgga cagttcccat 420
cccacagtca ggcggccatg aatttgtttg gaggcaacgc tttccaagga ggctgagtcc 480
atcgcccgat ggtgtggctg gtccggccgg ggcacagtgc agagctccta cccgggactc 540
tctctgacac ctagtgtggg agccaggcac actgcacaga cagacacatg gctgaggtat 600
gaccctccta gccaaccaaa aggcaagcag aggcgcacag gatgcaagca cgagaagagc 660
aacttgtcct cga 673

<210> 74
<211> 583
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (15)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (64)
<223> n equals a,t,g, or c

<400> 74
atamcatttc mcacnggaam crgctatgac cctgattacg ccagctcgaa ttaccctcac 60
taangggaac aaaagctgga gctccaccgc ggtggcggcc gctctagaac tagtggatcc 120
cccgggctgc aggaattcgg cacgagacag gtgcatgcac acgccactgt gtgtgtgtat 180
gtgtgtgtgt gtgtgtgtgt gtaggggaat cttagtctaa agcatcccac tgcaaactaa 240
aagctcttta aagtatatta atgtcacaaa aagttaaggc atttttccat tcttgttagc 300
atgtttcttt taccattttt ctcatttcaa attactttga ctttaaacgt tccctgaaac 360
ttaaatatac tgaggttctg ggaagagcta acatgccaac atttctattt tgatacacat 420
atctttctgg caagctgctg agtacctcca gttaagaagc acaggcctaa actctcagtg 480

tacagcattg ataaaatata tctcgagggg gggcccggta cccaattcgc cctatagtga 540
gtcgtattac aattcactgg ccgtcgtttt acaacgtcgt gac 583

<210> 75
<211> 801
<212> DNA
<213> Homo sapiens

<400> 75
gaattcggca cgaggatggg atttcaccat gttggccagg ctggtctcga actcctgacc 60
ttaggtgatc ggcctgcctc ggcctcacaa aatgctgaga ttacaggcgt gagcaccgca 120
cctggtctga tttttttttt aaatgcaaat cagactatgt cactcttttg cttgaagctc 180
ctcagtggct gcctatggct gtcagggtca gagcctcacc acggcctggg tttcctcctg 240
tggcccctgg ctttcgcctc ctgctctatt cttatcctga actacgccaa gccctttctc 300
aaccccgccc cttgctccct ctgtctggaa ctaccttccc aggccttttt gtgccgttca 360
ttctcaagtc acctcctcag cgagccctcc ttagtcactc cctttcatca ccctgtttgc 420
ttccttccca ttatctggtt tccctggagg cttatgtctg tctcccctca gtggaatgtg 480
ggcctcatgg cacaggccca tcggggtcac tgctgtgttc agggctcagt aaggatgcct 540
cggtgcgctt ggatgtggcg ctggccggct ggctgggggt gccacctggc gtgatttgtt 600
gtcacttgct cacttgtcct agatgctgtt tataaaagta ctaatagaac caggcacggt 660
ggtttatgcc tgtaatccca gcattttgga agcccaaggt aggcgaatcc cttgagccca 720
ggtgtttgag accagcttgg gcagcagggc aaaaccctgt ctctactaaa aaaaaaaaaa 780
aaaaaaaaaa aaaaaactcg a 801

<210> 76
<211> 982
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (554)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (615)
<223> n equals a,t,g, or c

<400> 76
gaattcggca cgagtggcca gctgggtaat gggaacagaa cagagcctgg gatatagggt 60
acagggctta ttattggtgt tatcacttca tgtctcccag agaggcctct gtgggtcact 120
gcccccctca atgagttctg aagagagaaa acagaggccg tggtccagtc agtatgggga 180
gcactgtgtt cccgacaccc cactgcgtgt taaggtcagg cgccacatct tgtagtcagt 240
tgctttgccg agtggctcca gctttctcta gctcctctct gggcctcagt ttccctgcct 300
gctggccaac agagggccct gccaactctg gctgcctatg accagggtgg ctccagaggg 360
tgctgctggg aggggtgcca accctamctc tctgcaagtg aaactgggca tgccamtcac 420
ctctctgggg cctcagtttc ctcttctgag cattgaggaa atttgggggt ttccatgttc 480
cttccagtca gaaaccagat gctgccatgt cccccaaccc aaggcctcag gaacagtgct 540
ggatggtcat tttngaggtt ttctgtctct gtctctccga ktgaggtttg cttggaaagc 600
taagaataga atccnagcma ggctgtaktg gcggccagct ggaacctgat ataktcacat 660
atgagaactg gtaggcctgc atgccgaccc tctatggacc agaatgggac agaggccaga 720
atatggccat gctcttcatc ctcactcctg ccccactgcc ctcagcccag tcctcctgtt 780

ccatctgact gaaaatcagg gcatgtaggg tgatggtttg ggctgcagcc aggtctgtgc 840
ctgctgtggc ctctgagctc tggagtcaaa tggggactgt ggaagaggct gcctagagtg 900
gcagaaaccc taccctggaa tggggagctg gctcagctgc ggctcactg tgtgagcctc 960
agcaacttgc catccctctc ga 982

<210> 77
<211> 1001
<212> DNA
<213> Homo sapiens

<400> 77
gaattcggca cgagtactct taagagcact caatcacctt ttgaatgctt tgctgcttag 60
aaatttcttc tgcaagacat cctaaaacat ctctctcaag ttcaaagttc cacagatctc 120
tggggcagag acaaaaatgc tgctagtctc tttgttaagc atagcaagaa ttacctttat 180
tttagttccc aacaagttcc tcatctccat ctgagatcac ctcagcctgg actttattgt 240
ccatatcact atcagcattt tggtcaaaat cattcaacaa gtcactaaga agttccaaac 300
tttcccacgt ctttctattt ttttctgagc cctccaaact gttccaacct ctgcctatta 360
cccagttcca aagttgcttc cacatttttg agtatyttat agcgsacccc accctctgca 420
gtmccawttt mccatattag tcmcttttcm cattactatg aagaaatmcc cagcctgggt 480
aatttattaa ggaaagatgt gcmattcact cacwtctctg cactaccagg gagatctcag 540
gaaacttaca tcmtggcaga aggcaagaag aagcagacac cttcttcaca gggtggcagg 600
acagagctag tgcaagtagg gaaaatgccc agatgcttat aaaatcatca twtctcatga 660
gaagtcactc actatcatga gaacagcatg agggaaacta cccccatgat ccaattgcct 720
ccatctggtt caccccttcat atgtggagat tatgaatatt accatttgag agagattttg 780
aatgggagca cagatccaaa ccatagcact gccttaaggt atctaataat caaactccca 840
aaggtcaagg gaaaagaaag gattctaaaa atagcaagag aaaagaaaca aattacatgc 900
aatggagcac caatatgtct ggctgcagat gtttcagtgg aaactttatt ggtttaggag 960
agagtggcat gacgtgctaa aaaaaaaaaa aaaaaactcg a 1001

<210> 78
<211> 748
<212> DNA
<213> Homo sapiens

<400> 78
tcgagggctg ggcctaactg gaactctgat tccatgtaga aaagacaggg tcccacagcc 60
tgcttccttt ctcctggctt ggtgggcatg cagaatttct tgacccctgt gttccaacaa 120
gagagctgaa aggaactctc ctaaagaact cacatatatt ttttaaattc taattctttt 180
tccaaataga agtttgaaaa ggcacccccct agaggaacat gcacttctgg actggcccca 240
ggttccagct tgggttggcg ggcgtgccag cagctcagtt tgaaacctct cacattgaat 300
caagggccag aagcagggcc tgtgggaagt ttctaggctt ctgctcatcc agaactgtcc 360
cctcagcatg gtgtgaggct cttatggagc ctgcagtcat aggatatgag acaaaaagcc 420
ttcccatcca tggatgtcct ttcatacact ggcacagaac accgggaaca aatgaaggtg 480
actgagaaaa aaaagaggca gacccatttc ttctgcgggt tgttttgtca tccagatacg 540
cttactttgt gcttatagct gtatgatctt ttttcctcat ctctaatgca caggatttct 600
gcctcattac ccatacagct aaagcttaat attaactaaa tcagtggtga attcctttcc 660
tttcccaccc cgacactatc agcgacattt ttcataatgg ccagcagagg tcagtgtgag 720
aacatagaga ctacactcgt gccgaatc 748

<210> 79
<211> 586
<212> DNA

<213> Homo sapiens

<400> 79
gaattcggca cgagggacta ccaacaagtg ttgctggacg tccggcggtc attgcggcgg 60
ttccctcctg gtgagaagct ctcccggtcc tgccacattt ggaaagactg tatctgttcc 120
aggtcatacc atgtgaccta atatgctgga ccctgccgcc tcagggacct tcagagctct 180
ccttttgctg agtcatcctt ttcttgactg gtcactttca gaccccact gtgaaagcct 240
gaaccaaaaa taatttctcc tggcctagag gtggtgaatg agagaagagg tttttgtttt 300
tccttgaagc cacaaaaagg agttaataag gattgttaga gccatcagtc tggcattaaa 360
gagcagattg gtgtggaatt gggcaccaac aagaatgagt aatatcttaa ttaggtttaa 420
aaacgatggt accttgcgca tacatatgta agattcctta gagggaagag aggccattcc 480
ctgtttgtgt aagagtatat tccttaatta acaaattaag cagcaataga taaaaaaata 540
aataaataaa aacaaaacaa acaaaaaaaa aaaaaaaaaa actcga 586

<210> 80
<211> 546
<212> DNA
<213> Homo sapiens

<400> 80
tcgacscacg cgtccgaaaa tactttttaa gaaagaaaat gacagaagca acccaagtgt 60
ctactgatgg ataattaaat tatagtatat aaatacaatg gggccgggtg cagtggctca 120
gccttccaaa gtgctgggat tacaggcatg agccacaaca tccagcccct tttctctttt 180
cttacccttc tttcctattt tcttttccat tttctttccc tcccttcttc tttctttcct 240
aactattaag gagtagattg aattcaaggt ctttatgtgt gtcagttttt gttttccaac 300
aaatatttct taaaaaccaa ccattgaaac gtaatggtaa ccactggccc ctgtctccac 360
ctccacacct aagaagcccc aaatccagat gtgtccatta aaatcagtcc agatcttctt 420
taccaagcca ctagatgtca tattaatttc acagcagaat agggaagccc atgccggagc 480
tgaaaacctg caacaacaaa aaagcatcta aatactgcaa aaaaaaaaaa aaaaaaaagg 540
gcggcc 546

<210> 81
<211> 708
<212> DNA
<213> Homo sapiens

<400> 81
tcgagttttt tttttttttt ttttaaatta gtcaaacatt ttattataga gtatatattt 60
atatcaaaag cacaaaaacg tttattctga aaaccaggaa gattgtgatg ttacagaaga 120
agattcaata attccagtcc atttctaggg tactaagtgt ctgatcacct cagygaaaac 180
aagatacaaa tgaggccaag gtcacaggtc tggccacccct gagtcccttc gcactatttg 240
gtttctcaag ttgagacacg tattcccagt cccagttagc caccttccaa gtgtttgcta 300
ctagccttaa tgggtactta gccaaagact acacccaaat ataaccaaag cttatgttaa 360
gtcataagat taatccttca ataataagga tagcataatt ggctttgtta cctaattcta 420
cataaacaaa atcatcaaat atcctggcat aactgaaatg acttacagag gaagtagtaa 480
agcttggaag tattctatgg taactgagct gaaaaagggg aaatgccaaa tgttgtaaat 540
gccatcatta ccaataagag tcaccaaatt ctcagaaata ggtaattggc agctcaaggc 600
agttagcact acaagatttc tcttgccttt aaaaaaaaat catttttaag actccttttt 660
taaaaggcta catcaaaaaa taaaccaaaa taacctcgtg ccgaattc 708

<210> 82
<211> 824

<212> DNA
<213> Homo sapiens

<400> 82

```
gaattcggca cgaggagaaa tttttcattt ttgattttta aaccattaga gcagtagctg     60
agcctttcaa gtttctcagt caagaattag gctatgagta gggacagttt tcttctctgt    120
ttcattttta ttttgttcc cttagtgaca ttgcaggaat gctgctgaaa tctacaggaa     180
gtttttaga atttggctta caggagagct gtgctgaatt ttggactagt gcggatgaca     240
gcagtgcttc cgacgaaatc aggttggagt tgtgcttcct ttccccttcc acttcttatc    300
tcgtagtttc cttcctcatg gtgagatcct agaaggagcc ttgttcaaac caaattgtgt    360
tggcctggaa gaatttgggc agtagatgta aagggattta tttataactg ccttgtcttt    420
tcatgtgatt tcttagttat ggttttatgt gaaattttct ttgaagggga acttagaatt    480
tatttagtgt gataaaaata gtgccaactg gctgggcgcg gtygctcacg cctgtaatcc    540
cagtactttg ggaggccgag gtgggtgaat caccaggtca ggagttcaag accagcctgg    600
ccaagatggt gaaacctcgt ctctactaaa aatacaaaaa aaacagctgg gcgtggtggc    660
acgcacccgt gatcccagct attcaggagg ctgaggcaga aaatttcttg aacccaggag    720
gcagaggttg cagtgagcca agatcatgcc actgcactcc agcctgggtg acagagcaag    780
actccgtctc aaaaaaaaaa aaaaaaaaaa aaaaaaaaac tcga                     824
```

<210> 83
<211> 789
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (789)
<223> n equals a,t,g, or c

<400> 83

```
gaattcggca cgagcttgag tattagctgc gtggttaagc tctatcatct gggactgcag     60
ggcctaagtt taaaaccttg agtgaaatta ttctgcttct ttaggcctca ttattctgaa    120
agactggata ggagtggtat ttatcccaga cggtagcttt gaatttggat ggagataatg    180
tatgtaaagg cctctgcagt cacggtctcc agagatgagg ctcttactcc ctgtcttcca    240
gatcctcact ggaatgcacc ctttgcaaga cacctcctcc agcccagctg ttcctttctt    300
gaattcccat agcacttcac tggtatttct ttctagcact taacagttat gtgcctgaca    360
tgatggttaa aattttacct tccctttgag actctgagca cctctaggct agggaagggc    420
ttggtgcact ccgtgtcctc tatacttgtg ggtaccaaac cgagaagagg atcaatatca    480
cttgaggagc tttgaaaaat agattccttt gggaggccga ggtgggccaa tcacaggtc     540
aggagattga gaccatcctg gctaatgcag tgaagccccg tctctactaa aaatacaaag    600
gattggctgg ccttggtggc gggcacctgt ggtcccagct acttgggagg ctgaggcagg    660
agagtggcgt gaacctggga ggcggagctt gcagtgagcc gggattgcgc cgctgtactc    720
cagcctgggc aacagagcga gactccatct caaaaaaaaa aaaaaaaaat cgaggggggt    780
cccgtaccn                                                            789
```

<210> 84
<211> 811
<212> DNA
<213> Homo sapiens

<400> 84

```
gaattcggca cgaggggcga tcatgctgag cgagactcca catgccagga gggggagggc     60
atttctcacc gacagtcttc ccatggtcat tccatccctc ctcctgcctc ctccaggcag    120
```

agcctctctg gctgagccca ctcttagatc tgtgaaaggg cagcctctca ccctgtcaca 180
gcacatggaa gaccttgctg tgagcagaga gaactgctcc cactataggg tccagctttg 240
tcctccagcc cctgcccctt cagctccacg ccttaccctg atggctcttt cctgctccag 300
cctcccctga gctgcccctt tcatcctatc tgcccctca actaatgcag cacagtctca 360
gtaaggtgat ctgtaactct ggctcaggg cttctcaggg ggactgaaga gtaacatcac 420
atcccatgaa cccactcagg gaggggcggg gctggtcatc actgagtcct cacttgaaag 480
aaagctgaac ttaggccggt tgtgctgggc acggtggctc acgcctataa tcccaacact 540
ttgggaggcc gaggcaggtg ggtcacctga ggtcaggaat tcgagaccag cctggccaac 600
atggtgaaac taaaaataca aaaaaattag ccgagcatgg tggcaggcac ctgtgatccc 660
agctactcag gagaatcgct tgaacccgga aggtggaggt tgcagtaagc cgagatcaca 720
ccactgcact ccagcctggg cgacagagcg agactccatc tcaaaaaaaa aaaaaaaaaa 780
ctcgaggggg ggcccgtacc caatcgccta t 811

<210> 85
<211> 1070
<212> DNA
<213> Homo sapiens

<400> 85
gaattcggca cgaggtgata cttctgaaga ctgcagggag aatccgtttt ccagcttttt 60
tcatccacca gaggccacct gtattcccta tcccacaacc ctagcccctt cctctatctt 120
tgaagtggac tatttcatcc cctgtttcta tcatgacagt gccttctctc atattgaccc 180
tcttgcctta taagattcct tgtgattaca ctgggtccac ctgcataatc aaggctaatc 240
tctccatctg gagatcttaa tataatcaca tctacaaagt ccctttggcc attgaagtaa 300
catatttata tgtattcatt attaggatgt gggacacttt tgtcagggac agggattttt 360
cagcctacct ttttcttcac cttttgccac cactctcagc ctgtggtctc aatgccagcc 420
tttacactgc taccccccatt gtctgggtag ktcataccag ycctcaagac tagcctcagg 480
cattgcctct tctgggaata catcctctta caggccagga tatgactcat gggtgcattc 540
ctaatagcac ttcamttatt tctactgtca ccacactgat ctgtaattac ttgatttgtc 600
tgactcttct gggggcttgt aagcattctg gcacagagaa ctatgactta ctggggctta 660
catctcttgc taaacacagt acctaaaatt tagtaggcat tccctcataa acatgaatga 720
atgaatcaaa gaatgaataa acatttagga aatgatgttg tgttggtcaa cttctttcct 780
catcactgtt aaagataaaa gaatgccaag ccaggttgtt cagacagaag caagcaccac 840
atccctgaga gagcagcaca tctgggcagc catgtgtgag aagtcggttg cattccccat 900
acacagttgt ctttgcagct gtactcttaa ccactgtaac cacagaagtg ggaaacaat 960
agggtggggt gaagtgaaaa gaaaattttc caaaacttca tttatctaat aaatacagat 1020
atttaaaaaa aaaaaaaaaa aactcgaggg ggggcccgta cccaatcgcc 1070

<210> 86
<211> 727
<212> DNA
<213> Homo sapiens

<400> 86
gaattcggca cgagagggtt ttagtttatg tctctaactt tagcaaagct gcattcctat 60
tggaatgcat actggaaaca gctctcattc ctacctttaa agggctcttg gaaagcagtg 120
tgacaaccaa ggtcactaaa tggtgagatc atcaagccat tttaagttct ttctcatgtt 180
attcaccagc accctgcagg acgttgggca cacatcacat ccctcagctc agccatccag 240
ccgtctcagt gattcaccac tcatttgctt aattaataga caggtttgat cactttgtac 300
atggaaggca ctgtgccagt gaacaagcag ttggacccag ccctccagta gggaatggac 360
agctgaaaat ccatgagcaa gaaagaagga aaaagaaaga gttctgagca gccaaaccat 420
ttctcgatga tttcagagcc ttcattctga gcatcagtta tatgctctcc agtgtaatga 480
ctttatagcc aagcacagta attgatatta ctgtgaaggc ccttaactta tcaagaaatg 540

gttgaggccg ggcacattgg ctcatgccta taatcccagc acgtgggagg ccgaggcagg 600
cagatcactt aagcccagga gttcaagccc agcctgggca acatgatgaa agcccatctc 660
tacaaaaaaa aaaaaaaaaa actcgagggg gggcccggta cccaattcgc cctatagtga 720
gtcgtat 727

<210> 87
<211> 690
<212> DNA
<213> Homo sapiens

<400> 87
gaattcggca cgagagcagg gctaggtgga catgaggagc ccagttcagg gctgtcacag 60
tagctccagc agcagatgat tgtggctggg cctcccaagt gtcacgttgg agaaccggag 120
aaggggactt ctttgggatg tactctggac ttgttgatag attaagtgta ggtggggtga 180
ggaagagaac tcaaagatga caccaggtgt tggagctgag ccacggggag aagggtgcaa 240
agggaaagca gtgcgggggc tgggagggga gagggtcagt cctgttttgc ttgtgctgca 300
tctgaggagc ccctcacctg tggaaggaga gcagtcccag aggcagtggg gtgtgcagtt 360
ctggaactta gaagaatgat caggggggctg ggtgcagtgg ctcacgcctg taatcccagc 420
actttgggag gccgaggcgg gcggatcaag aggtcaggag attgagacca tcctggctaa 480
catggtgaaa ccccgtctct actaaaaata taaaaaatta gcagcgcatg gtggcaggca 540
cctgtagtcc cagctattca ggaggctgag gcaggagagt ggcgtgaacc cgggagacgg 600
agcttgcagt gagctgagat tgcgccactg cactccagcc tgggcgacag agcgaactcc 660
gtctcaaaaa aaaaaaaaaa aaaaactcga 690

<210> 88
<211> 896
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (401)
<223> n equals a,t,g, or c

<400> 88
gaattcggca cgagaaattg agaaacatta atacaaagta agagacaaga gcctagtaac 60
aaatggtggc tctttgagaa aaggaaatta ttaccaaaat tttagactaa ctgaaggcat 120
gccaattaag caccagattt tgctcttaaa ctttttggga agctgagtag aaattatcct 180
tttgttccat atgatgactt attaaataaa atactttgca caatatgtgc ttttagatgg 240
agtaaacaac ataccttttа aataattatt ttgattgcct atattcatat catgatgcta 300
cctttkgca tttgtgcagt gtacatkgaa tattaactga gtgtttagaa atgctggatt 360
ttaggtttca gctttgctgt gggtgaaggg aagtgggggg ncttctgttt gttggtgcca 420
ggcattatgc tacatattat acatctgtta tctcatttga ttycccсaaa tccttaagaa 480
gttgaattat tatactcatt ttggaaataa gaaatgaagc ttagagaggg gaagaacagg 540
tttaaatcct ggctgtaagc cctttgggct ttggttttcc taactaggga agaggaataa 600
tagtgatgaa aataacaatc atctgatgat ctttgtaatt ttactgacgg agtagaagcc 660
atcagaagag aatgcccaca tcttcccttt gatagagcat ctgacttgca tctccttagt 720
aactactttc cctcccattc taaactgttc ttttctaggg gccaacctct cctcttgtga 780
acgagctctc atcctttcct ggatacacag cttcttcttt cctgcatact tttttctttg 840
tacagcatga aaatatacta ttgtgtcttg tttaaaaaaa aaaaaaaaaa actcga 896

<210> 89

<211> 857
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (550)
<223> n equals a,t,g, or c

<400> 89
gaattcggca cgagcttatg gtctttatta cttacatgtc tcatcattcc tgtacaactg 60
tagcaaatat aaacatcaaa atgcctctag atcttctttt cctcataaca tattttctcc 120
tttctgtcat tttgaaagtg ttgtatattg atgcccctgg tcatttagga atgcccattt 180
ctctttgttc tagtgctgtt gtgtgggtga aggttgacct agtktcagag aaggggtgag 240
gaaaggcagg ggcmaaaaga ataaaggaaa gagttycttt tgagtacmaa taaaaactac 300
caggaaaatc tgatttacca aaatgttcta gggattagat tgcaacyatt aaatatgatt 360
taacygaagg acccctccgg cctttttat tcccttcttt tttactaaaa ttctttatcg 420
aattgcagaa tccttttca ttkgtctcag taagtaaact tcaataaatt ataggtaaaa 480
tttagaaaac tgaaaattct gttagagatt agaatgcatt aatatttctt gccttaggct 540
gggtgcagtn gctcacgcct gtgaccccag cactttggga ggctgaggcg ggcacatcac 600
ctgaggtcag gagttcggga ccagcctggc cgacgtggtg gaaccccgtc tctactagaa 660
atacaaaagt tggccaggca tggtggcagg cccggctact tggtaggctg aggcaggaga 720
atcgcttgag ccagggaggt ggaggttgca gtgagccgag atcgtgccac agccgagatc 780
tgtgagcctg ggccacagag cgagactcca tctcaaacaa acaaacaaac aaacaaaaaa 840
aaaaaaaaaa aactcga 857

<210> 90
<211> 561
<212> DNA
<213> Homo sapiens

<400> 90
agggatcccc cgggctgcag gaattcggca cgagtctact ctcaaaaaat tcagaaacat 60
atatttgtgt gcatttgcat gtgcaacagt acacacaaac atacataaag agagcaattg 120
ataaggcaaa taaggtaaca tttaacaata atctgataca cataaataga gaaagagcaa 180
ttgataaagt aaatgaggta aaatttaaca ataatctgag caaaaggtat atgtgttttc 240
tttgagacag tctgattctt gcaacttatt ctgtaagttg gaacttattt ccaaacatga 300
ttgaaaaaaa accccgcact tggcaacttc ttctcttttt cagcctagaa atgtctgtgt 360
taagtggttt tttatttatt gttgttgttt gttgttattg ttgttttgtt gccaggctcc 420
aactcacaaa atacgagttt aaaaactgcg ttgttatttt tagagatttg tgataataca 480
acttgttata aaatttattc ctcaataaat ataatttctc tactaaaaaa aaaaaaaaaa 540
aaaaaaaaaa aaaaaactcg a 561

<210> 91
<211> 655
<212> DNA
<213> Homo sapiens

<400> 91
gaattcggca cgagctcaaa caaacaaaca aacaaacaaa ctagcatgga gagggacaca 60
agagagaaat gtttatggtc cttgccttac cctaaattac tgtgcaacct tttggcaagt 120
cacttcctct ctattctgag tttctttatc tattcaattg ggttcttaga tttggtggtc 180
tctaacactc tcccagtttt tcaatttgat gttacattct acccagtgac caaattcata 240

ttccagaagc atagtatgct atgtcatacc gcaaatcttg taaacgttcc tgatatggtt 300
tggctgtgtc cccacccaaa tctcatcttg aattgcagtt cccataatcc acacatgtaa 360
caggagggac caggtggagc taattgaacc atgggggcga tctcccccca cctgttcttg 420
tgatagtgag ttagttctca tgagatctga tggttttata agggtctttc cccttcactg 480
ggcactcatt cttctgcctc ctgttgccac atgaggaagg acatgtttgc ttccccttct 540
gccatgattg taagtttcct gaggcctccc agctatgctg aactgagagt caattaaact 600
tttttccttt ataaatttaa aaaaaaaaaa aaaaaactct gacggggggg ccctg 655

<210> 92
<211> 848
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (2)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (17)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (81)
<223> n equals a,t,g, or c

<400> 92
cnaggccwrr aaccccnaag gctggcactg agctgtgact gctttaacag cccccaagat 60
ttggtcagtt tgaggtggtg nagactcaga ttgttgctga aagttcagta acacagtcct 120
ggtctttggc cctagagaaa ctttttatat gagaagtgtt ctctatatac atgtttgagg 180
tgactctgga atggattatg aggtcatatc tcaaaatgtc agaaaacgtt atagagcact 240
cgaacttttg tatttgctgc ttaacctcaa tattacagcc acaaacaagg ggtaccaaga 300
caaagtataa ctgagcataa gcagaaaatg ttaaccctcc aggtttcttt cttaagcaca 360
ataaaagtgg gagcgaacaa cacaaggata tttttacatt tgacccgtct caaaagtagc 420
acaccctatc cttgtgccat tatttgtaca aggaaatata tgattagaag gawtagaacc 480
cccagttgtc atcagctttt ttagacacca caggttgtag cagtttgaac aaactgaaaa 540
ctttatactt ctgtgtgagc tgaactcaag tttcagaata atcatcgcca tgtgggaggc 600
tttttgttaa atgcagaaga aatttcaaaa tattgtattt atatctgcct tccactgctg 660
ccaatttagt aagcatctcc tatacaatcg acaataaaca gcaaatgatg cagttcatag 720
agtattttgc acttggggaa aaatatgtat ctgaattgta aaaagaaatg tttggatttt 780
gtatgtcttt tttattatta ttaaaatact aaatgaaact cctcaaaaaa aaaaaaaaaa 840
aaactcga 848

<210> 93
<211> 612
<212> DNA
<213> Homo sapiens

<400> 93
gaattcggca cgagagcgtg ttattctcct gcctccagat catttaggct ttggtaaaac 60
ctcggccaat ttggctataa taaaatagat ttccttgagg gcaggattgg ttaggggaa 120

```
cagaaagctc tgggtattat ttcaaaatga tttattttct cctcctcttg cctgaagcac     180
aaggagagtt ctcatcgatt ttcacagtga gaacctggta ggtaatactc atttaagcat     240
gggatcctgt gttcgtccag acccttggag ttttaaattc tcagggtggt tcaacctgag     300
ttaatttgtc aattatggtt taaagtgttc ctatggatgt tggctttagc tgcaggctcc     360
tgtatccatc tccctctcta gtttttgaga tggcagtttg tttcatgacc tctatgaaga     420
gctgccatct atctatctat ctatctatct atctatctat ctatctatct atatacctat     480
ctacctatct atgagaggag tcttccttga gcccaggagt tcaaggttgc agtgagccat     540
gatcatgcca ctacactcca ccctcagcaa cagagaaaga cactatctca aaaaaaaaaa     600
aaaaaaactc ga                                                          612
```

<210> 94
<211> 951
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (826)
<223> n equals a,t,g, or c

<400> 94

```
gttacctccc gcagccgcag ccgccgtgct cagcgcgagc cccggagccc ttgagcgcga      60
ggcgcggagc cccccggagcc cccaaaccgc agacacatcc ccgcgcccca gagcccggc     120
ctgcgcgccc agccgggccc gcgcgatgcc ctcagaccgg cctttcaagc agcggcggag     180
yttcgccgac cgctgtaagg aggtacagca gatccgcgac cagcacccca gcaaaatccc     240
ggtgatcatc gagcgctaca agggtgagaa gcagctgccc gtcctggaca agaccaagtt     300
tttggtcccg gaccatgtca acatgagcga gttggtcaag atcatccggc gccgcctgca     360
gctgaacccc acgcaggcct tcttcctgct ggtgaaccag cacagcatgg tgagtgtgtc     420
cacgcccatc gcggacatct acgagcagga gaaagacgag gacggcttcc tctatatggt     480
ctacgcctcc caggaaacct tcggcttctg agccagcagt aggggggctc ggcctgggag     540
tcgggggggcc ccggtcaggc cctgcccaga gagctcctgg ttcctgaact gagctgcctc     600
taccgtggtg ggctgggcag gcatgtgccc ccctagtcag agggcaccaa cccacctayt     660
ctgcccctgg gtggatcctg ggccggtcgt gttagggttg tccctctggg tgctggctgg     720
tgggatgggg gagggtgggg agcagctccc agcacccctg ctgtgtggtt catctttttt     780
ttaggcccct gcctgtctgc ccatctgccc ctcacccacc cgaggntctg cccaccgcct     840
ggacctgccc acccctgaaa gactggcccc tggctccccg cccctcggtc tccacgtggt     900
gtatggatct gtggtcattg tccctctgca gaataaagat tgctcaggcc t               951
```

<210> 95
<211> 2264
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (299)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (2257)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (2264)
<223> n equals a,t,g, or c

<400> 95
aaatttctca acaccacagt cagctaagtc acctactgcc accttcgaaa aacacggaga 60
gcacctaccc agaggagaag gtagatttgg agtaagccgc cgtcgacata attcctctga 120
tggttttttt aacaatggtc ccctacgaac tgcaggagat tcttggcacc agscctccct 180
gttccgccat gattctktgg actctggwgt ctctaaggga gcatatgctg gaatcacagg 240
gaacccatct ggttggcata gctcttcccg aggtcatgat ggcatgagcc aacgtakgna 300
ggtggcacag ggaaccatcg ccattggaat ggcagcttcc actcccggaa agggtgtgct 360
tttcaggaaa agccacctat ggagattagg gaagaaaaga aagaagacaa ggtggaaaag 420
ttgcagtttg aagaggagga ctttccttcc ttgaatccag aagctggcaa acagcatcag 480
ccatgcagac ctattgggac accttctgga gtatgggaaa acccgcctag tgccaagcaa 540
ccctccaaga tgctagttat caaaaaagtt tccaaagagg atcctgctgc tgccttctct 600
gctgcattca cctcaccagg atctcaccat gcaaatggga acaaattgtc atccgtggtt 660
ccaagtgtct ataagaacct ggttcctaag cctgtaccac ctccttccaa gcctaatgca 720
tggaaagcta acaggatgga gcacaagtca ggatcccttt cctctagccg ggagtctgct 780
tttaccagtc caatctctgt taccaaacca gtggtactgg ctagtggtgc agctctgagt 840
tctcccaaag agagtccctc cagcaccacc cctccaattg agatcagctc ctctcgtctg 900
accaagttga cccgccgaac caccgacagg aagagtgagt tcctgaaaac tctgaaggat 960
gaccggaatg gagacttctc agagaataga gactgtgaca agctggaaga tttggaggac 1020
aacagcacac ctgaaccaaa ggaaaatggg gaggaaggct gtcatcaaaa tggtcttgcc 1080
ctccctgtag tggaagaagg ggaggttctc tcacactctc tagaagcaga gcacaggtta 1140
ttgaaagcta tgggttggca ggaatatcct gaaaatgatg agaattgcct tcccctcaca 1200
gaggatgagc tcaaagagtt ccacatgaag acagagcagc tgagaagaaa tggctttgga 1260
aagaatggct tcttgcagag ccgcagttcc agtctgttct ccccttggag aagcacttgc 1320
aaagcagagt ttgaggactc agacaccgaa accagtagca gtgaaacatc agatgacgat 1380
gcctggaagt aggcatataa atgctcacag ttaaatctga cccagtaaac tctgtgtgtt 1440
tagggagtat acaaaagaaa tcgttctttt ccttttctta tgttgttgaa tacttcattc 1500
acaagggaaa taatcatatc ccaaagagag agcaattggc ttgttttgct tttgttattg 1560
ttcttccctg ttatctgctt tatagagaga agtttgtgtg gtgggacaga ttttttaaac 1620
acactcayac acacacacac atacacaccc agtatatatg gggcgatgca caggtaggag 1680
ctggcagtgc agggaagagg agacactggt ctgcagcaac agcttctact accagccctt 1740
ggggcactca cccctgtgat caagcaatca ttgtcaatga caaagtgact attgaagtta 1800
taattgtatt aaattaatgc taataatttg gatattttat tttatttttg gctgctcggg 1860
taactttagc ccttaaccaa gcatatgtgg gttttttttgg ttgttttttt ttgttttttt 1920
tttctttttc ctttttgggt acagctgtaa aatatttgga tataggaaat gttgtgttat 1980
tcttgcagcc ttgatattca gggtggattg taaaatataa atttttgtga gatttcaaag 2040
attaagatta ttttgataac attatttaca gatttaaaag atgtggttat cacaagtctc 2100
gaggggggaaa ctactgcata aaataactaa cttggaataa atattttgca tcagtttgga 2160
taaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2220
aaaaaaaaaa aaaaaaaaaa aaaaaaaagg ggggggnccc cccn 2264

<210> 96
<211> 830
<212> DNA
<213> Homo sapiens

<400> 96
ggtcgaccca cgcgtccgct gaaaggaaaa gcactgtttg gagaatgatc caccttttcaa 60
gattttactt attgttgata atgctcccac atgtcctctt ttttacgggt gatcttcatt 120
cctaatatca aagtgatatt tcttcctcca ggcaccacct ctttgatcca cacaatggat 180

caaggagtta tagcagcttt taagttctac tacctgagaa gggaggactt ttgcccagtc 240
ccatactgca gtggaggaag acactgagaa gactctgatg aaattctgaa cagcatcaag 300
aaccttgttt aggcttggat tatgtcgcta aggactgtag gaatggcacc tggaagaaga 360
cacgcaagag gtttgtcaat aacttcaaag gatttgccaa ggatgaggaa gttgcaaaaa 420
tcaagaaggc tgtggttgag atggcaaaca actttaacct gggtgtggat gtggatgaca 480
ttgagtaatt cctagagggg gttcctgagg aattgactaa tgggttgctg ttggaactgg 540
aataggagtg catagctgaa gaagaggtaa agaaaaagaa agtgcaggag aagggaaaaa 600
agaactccca agaatactca cagtgatggg tttagcagaa gcttcttcag actccaacaa 660
gctccttaag aagtctgaaa acatggaccc caaaactgaa aggttttcac taatagagag 720
gaaagttcat ggtgcattat ctgcctacaa gcaaaaccag gattcaaaaa accctttgag 780
ctggagcttc aaagcacaaa aaaaaaaaaa aaaaaaaaaa aagggcggcc 830

<210> 97
<211> 886
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (92)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (886)
<223> n equals a,t,g, or c

<400> 97
atttcatttt agggcatact gggcttactc tcctcccagc tgtctgtgga ttgatttgat 60
tttaatgttc gagttttaca gcaacagctg anaaaccatg aactattcta ggaactgtgt 120
tggaactctt taaaataaag aaaagaggag gaggagagga agaaagaaaa ccaacttaag 180
aagccttgga ctttggaggg acagaaagcc accagccaat ggagaacaaa gagatgtttc 240
cctttccttt ctttcacctt gtcattctgg gtttccttct gcttcactct ttccttcccc 300
cttaaaagtg gtattcctgg ttggtctgtc tgtctgtcct tgtccttgtg gtgatcctgg 360
catggtgata tgctccactt tgcattatcc atggtctctt accagcgcac aagtcagtgg 420
ggaggatcta accacgcctg gtggtgagga agctgaattt ccaggcctgc gtcccatgta 480
gcctctccat gaactgcaga aggcatgttc tgcatggtta ccagtaagtg gctccctctc 540
accgtgttca ttgtcaaatg agagcaaact ttaggtgttg gctccattgt acactctact 600
tgctctgctc ccctccctcc aaccagggtt catgtcagtg cacaccccat gtgccctggc 660
gaagctggtg ctgtgagtga tgtttcccat acaactcagg gatgccaggt ggcttaccct 720
gagatagtca ttttgggcac ataacagtgt aggaatgaaa catggatttc attgatattt 780
aaatctgtca atttcatttt ttgttaatgt tttcccctga tgacttttta gcaatttaac 840
aaataaaatg gacaattgtc ttaaaaaaaa aaaaaaaaaa ctcgan 886

<210> 98
<211> 597
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (12)
<223> n equals a,t,g, or c

<400> 98
cttgccatat ancaagctga attacctcat aaggaacaaa gtggagytca cgcgktgcgc 60
cgtctagact atgatccccg gctgcagaat tcggcacgag cagtccagaa actgcgtgcc 120
ctgccctttg cttgggcccc tctaccagta tgtccagcat gtgcccgggg gccctcagct 180
cccctggggc ccagcccacc caagacacag ctcttggtcg tgaacatgaa gatgagccaa 240
actctagtgg ctcttcctga aagaaatgag aatgcccagc cacacccatg cacgctttgt 300
tctttttat ttaatactga ggaaccggag tggaggggtc ctgccgggct gcagtgaccc 360
tgagggaagt caggagagcc ctgggctgca gaagagtccc cccacaggct ccgaagcaag 420
cttgtcctgg tgcattcaga ctgctcacag caggctttgg gccctcactc tccagatccc 480
agagagccct ccagggctcc cagctctcgg gccagtgccc amgtcctcga agggggccg 540
gtaaccaatt cgccctatag tgagtcgtat tacaattcac tggccgtcgt tttacaa 597

<210> 99
<211> 66
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (66)
<223> Xaa equals stop translation

<400> 99
Met Phe Leu Gly Asn Ser Leu Glu Thr Leu Thr Asn Arg Ile Leu Val
1 5 10 15

Ser Leu Ala Ser Val Phe Leu Leu Pro Pro Arg Lys Gly Ala Gly Leu
20 25 30

Cys Ser Arg Gln Asp Arg Arg Ala Pro His Ala Tyr Thr Ser Leu Pro
35 40 45

Glu Leu Ser Pro Arg Ala Ser Gly Pro Cys Leu Glu Thr Gly Leu Ala
50 55 60

Leu Xaa
65

<210> 100
<211> 72
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (72)
<223> Xaa equals stop translation

<400> 100
Met Tyr Gln Glu Thr Arg Ser Ser Pro Thr Asn Thr Leu Arg Pro Trp
1 5 10 15

Pro Arg Gly Thr Ser Arg Cys Leu Arg Cys Ser Phe Cys Arg Leu Ser
20 25 30

Phe Ala His Ser Gln Gly Ile Gln Gln Leu Ser Cys Ser Leu Ser Arg
35 40 45

Thr Asp Ser Arg Ser Phe Thr Ile Ser Lys Thr Leu Trp Ala His Asn
50 55 60

Arg Arg His Ser Phe Gln Gly Xaa
65 70

<210> 101
<211> 51
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (51)
<223> Xaa equals stop translation

<400> 101
Met Asn Ala Tyr Arg Val Lys Pro Ala Val Phe Asp Leu Leu Leu Ala
1 5 10 15

Val Gly Ile Ala Ala Tyr Leu Gly Met Ala Tyr Val Ala Val Gln His
20 25 30

Phe Ser Leu Leu Tyr Lys Thr Val Gln Arg Leu Leu Val Lys Ala Lys
35 40 45

Thr Gln Xaa
50

<210> 102
<211> 221
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (221)
<223> Xaa equals stop translation

<400> 102
Met Asn Val Phe Arg Ile Leu Gly Asp Leu Ser His Leu Leu Ala Met
1 5 10 15

Ile Leu Leu Leu Gly Lys Ile Trp Arg Ser Lys Cys Cys Lys Gly Ile
20 25 30

Ser Gly Lys Ser Gln Ile Leu Phe Ala Leu Val Phe Thr Thr Arg Tyr

35 40 45

Leu Asp Leu Phe Thr Asn Phe Ile Ser Ile Tyr Asn Thr Val Met Lys
50 55 60

Val Val Phe Leu Leu Cys Ala Tyr Val Thr Val Tyr Met Ile Tyr Gly
65 70 75 80

Lys Phe Arg Lys Thr Phe Asp Ser Glu Asn Asp Thr Phe Arg Leu Glu
85 90 95

Phe Leu Leu Val Pro Val Ile Gly Leu Ser Phe Leu Glu Asn Tyr Ser
100 105 110

Phe Thr Leu Leu Glu Ile Leu Trp Thr Phe Ser Ile Tyr Leu Glu Ser
115 120 125

Val Ala Ile Leu Pro Gln Leu Phe Met Ile Ser Lys Thr Gly Glu Ala
130 135 140

Glu Thr Ile Thr Thr His Tyr Leu Phe Phe Leu Gly Leu Tyr Arg Ala
145 150 155 160

Leu Tyr Leu Ala Asn Trp Ile Arg Arg Tyr Gln Thr Glu Asn Phe Tyr
165 170 175

Asp Gln Ile Ala Val Val Ser Gly Val Val Gln Thr Ile Phe Tyr Cys
180 185 190

Asp Phe Phe Tyr Leu Tyr Gly Thr Lys Gly Arg Ser Trp Asp Asp Ser
195 200 205

Asn Ala Asp Thr Gly Leu Arg Ser Tyr Ser Ser Ile Xaa
210 215 220

<210> 103
<211> 114
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (114)
<223> Xaa equals stop translation

<400> 103
Met Leu Ser His Val Phe Pro Ile Cys Thr Arg Pro Cys Leu Ser Met
1 5 10 15

Tyr Phe Pro Cys Val Pro Ser Met Tyr Leu Val Tyr Phe Leu Pro Leu
20 25 30

Asn His Gly Ile Leu Leu Thr Glu Pro Tyr Val Pro Tyr Pro Ala His
35 40 45

Cys Tyr Ala Leu Phe Pro Asn Ser Cys Leu Val Gly Pro Ser Thr Pro
50 55 60

Ser Pro Cys His Arg Ile Ser Ile Ser Ala Gln Ile Pro Pro Ile Ser
65 70 75 80

Ile Ala Phe Met Tyr Tyr Pro Gln Ser Thr Leu Thr Ile Ile Phe Ser
85 90 95

Gln Asp Cys Ser Leu Leu Phe Cys Val Phe Leu Arg Gly Ile Lys Glu
100 105 110

Lys Xaa

<210> 104
<211> 132
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (6)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (132)
<223> Xaa equals stop translation

<400> 104
Met Glu Asn Ile Ser Xaa Asp Val Ile Val Gly Arg Cys Leu Ala Ile
1 5 10 15

Leu Lys Gly Ile Phe Gly Ser Ser Ala Val Pro Gln Pro Lys Glu Thr
20 25 30

Val Val Ser Arg Trp Arg Ala Asp Pro Tyr Val Ala Ala Gly Ser Ser
35 40 45

Gly Asn Asp Tyr Asp Leu Met Ala Gln Pro Ile Thr Pro Gly Pro Ser
50 55 60

Ile Pro Gly Ala Pro Gln Pro Ile Pro Arg Leu Phe Phe Ala Gly Glu
65 70 75 80

His Thr Ile Arg Asn Tyr Pro Ala Thr Val His Gly Ala Leu Leu Ser
85 90 95

Gly Leu Arg Glu Ala Gly Arg Ile Ala Asp Gln Phe Leu Gly Ala Met
100 105 110

Tyr Thr Leu Pro Arg Gln Ala Thr Pro Gly Val Pro Ala Gln Gln Ser

115 120 125

Pro Ser Met Xaa
130

<210> 105
<211> 88
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (88)
<223> Xaa equals stop translation

<400> 105
Met Glu Asn Thr Phe Phe Val Phe Leu Val Ser Ala Leu Leu Leu Ala
1 5 10 15

Val Ile Tyr Leu Asn Ile Gln Val Val Arg Gly Gln Arg Lys Val Ile
20 25 30

Cys Leu Leu Lys Glu Gln Ile Ser Asn Glu Gly Glu Asp Lys Ile Phe
35 40 45

Leu Ile Asn Lys Leu His Ser Ile Tyr Glu Arg Lys Glu Arg Glu Glu
50 55 60

Arg Ser Arg Val Gly Thr Thr Glu Glu Ala Ala Ala Pro Pro Ala Leu
65 70 75 80

Leu Thr Asp Glu Gln Asp Ala Xaa
85

<210> 106
<211> 64
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (64)
<223> Xaa equals stop translation

<400> 106
Met Ser Ala Ala Ser Phe Trp Pro Arg Pro Val Ala Ser Ile Ser Val
1 5 10 15

Phe Ile Leu Leu Gly Ser Ser Val Thr Thr Ser Lys Thr Arg Ser Gly
20 25 30

Val Ile Ser Ser Ala Gly Lys Pro Ile Trp Val Gln Ser Pro His Leu
35 40 45

Ala Leu Leu Glu Val Leu Leu Gln Lys Gly Ile Val Pro Glu Lys Xaa
50 55 60

<210> 107
<211> 41
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (41)
<223> Xaa equals stop translation

<400> 107
Met Leu Ser Leu Thr Val Ser Leu Lys Ser Val Ser Ile Ala Ala Gln
1 5 10 15

Ser Leu Phe Leu Asp Leu His Phe Pro Ile Gln Met Thr Leu Val His
20 25 30

Lys Glu Ile Ala Lys Leu Glu Thr Xaa
35 40

<210> 108
<211> 48
<212> PRT
<213> Homo sapiens

<400> 108
Met Thr Leu Tyr Leu Asn Thr Asn Lys Asn Lys Pro Ser Ala Leu Tyr
1 5 10 15

Ser Leu Phe Phe Cys Phe Ile Ser Thr Pro Tyr Thr Tyr Gly Leu Gln
20 25 30

Ile Cys Tyr Lys Cys Phe Phe Ile Tyr Ile Phe Val Ile Cys Leu Tyr
35 40 45

<210> 109
<211> 38
<212> PRT
<213> Homo sapiens

<400> 109
Met Phe Leu Thr Tyr Leu Thr Tyr Asn Val Ile Ser Leu Asn Glu Val

1 5 10 15

Val Ser Thr Ser Ala His Gln Ile Ala Val Ile Val Asn Tyr Leu Phe
20 25 30

Met Gly Asp Asn Leu Phe
35

<210> 110
<211> 45
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (45)
<223> Xaa equals stop translation

<400> 110
Met Pro His Pro Ile Trp Cys Tyr Arg Asn Ser Ala Arg Lys Val His
1 5 10 15

Leu Phe Ala Cys Leu Phe Ile Leu Tyr Ile Leu Pro Ile Leu Tyr Ser
20 25 30

Cys Thr Lys Asp Leu Ile Glu Asn Leu Lys Ser Ser Xaa
35 40 45

<210> 111
<211> 39
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (39)
<223> Xaa equals stop translation

<400> 111
Met Leu Arg Ile Lys Ser Cys Leu Leu Leu Phe Phe Ile Phe Phe Pro
1 5 10 15

Phe Asn Ile Lys Asp Ser Gln Val Pro Ala Asn Tyr Ile Ala Thr Phe
20 25 30

Ser Arg Lys Cys Ser Phe Xaa
35

<210> 112
<211> 25
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (25)
<223> Xaa equals stop translation

<400> 112
Met Ser Leu Gln Pro Pro Phe Val Met Leu Leu Leu Ser Thr Ala Gln
1 5 10 15

His His Glu Leu Gly Ala Asp Thr Xaa
20 25

<210> 113
<211> 50
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (50)
<223> Xaa equals stop translation

<400> 113
Met Pro Lys Gly Ile Leu Val Ser Phe Leu Cys Ala Leu Ser Pro Arg
1 5 10 15

Thr Gly Met Leu Gly Val Ser Phe Leu Leu Phe Ile Gly Ile Leu Leu
20 25 30

Arg His Thr Ser Cys Leu Phe Cys Met Val Phe Ala Lys Met Pro Leu
35 40 45

Ala Xaa
50

<210> 114
<211> 54
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (54)
<223> Xaa equals stop translation

<400> 114
Met Cys Pro Pro Ser Gln Arg Ala Pro Thr His Leu Leu Cys Pro Trp
1 5 10 15

Val Asp Pro Gly Pro Val Val Leu Gly Leu Ser Leu Trp Val Leu Ala
20 25 30

Gly Gly Met Gly Glu Gly Gly Glu Gln Leu Pro Ala Pro Leu Leu Cys
35 40 45

Gly Ser Ser Phe Phe Xaa
50

<210> 115
<211> 268
<212> PRT
<213> Homo sapiens

<400> 115
Met Glu Val Ala Glu Pro Ser Ser Pro Thr Glu Glu Glu Glu Glu Glu
1 5 10 15

Glu Glu His Ser Ala Glu Pro Arg Pro Arg Thr Arg Ser Asn Pro Glu
20 25 30

Gly Ala Glu Asp Arg Ala Val Gly Ala Gln Ala Ser Val Gly Ser Arg
35 40 45

Ser Glu Gly Glu Gly Glu Ala Ala Ser Ala Asp Asp Gly Ser Leu Asn
50 55 60

Thr Ser Gly Ala Gly Pro Lys Ser Trp Gln Val Pro Pro Pro Ala Pro
65 70 75 80

Glu Val Gln Ile Arg Thr Pro Arg Val Asn Cys Pro Glu Lys Val Ile
85 90 95

Ile Cys Leu Asp Leu Ser Glu Glu Met Ser Leu Pro Lys Leu Glu Ser
100 105 110

Phe Asn Gly Ser Lys Thr Asn Ala Leu Asn Val Ser Gln Lys Met Ile
115 120 125

Glu Met Phe Val Arg Thr Lys His Lys Ile Asp Lys Ser His Glu Phe
130 135 140

Ala Leu Val Val Val Asn Asp Asp Thr Ala Trp Leu Ser Gly Leu Thr
145 150 155 160

Ser Asp Pro Arg Glu Leu Cys Ser Cys Leu Tyr Asp Leu Glu Thr Ala
165 170 175

Ser Cys Ser Thr Phe Asn Leu Glu Gly Leu Phe Ser Leu Ile Gln Gln
180 185 190

Lys Thr Glu Leu Pro Val Thr Glu Asn Val Gln Thr Ile Pro Pro Pro
195 200 205

Tyr Val Val Arg Thr Ile Leu Val Tyr Ser Arg Pro Pro Cys Gln Pro
210 215 220

Gln Phe Ser Leu Thr Glu Pro Met Lys Lys Met Phe Gln Cys Pro Tyr
225 230 235 240

Phe Phe Phe Asp Val Val Tyr Ile His Asn Gly Thr Glu Glu Lys Glu
245 250 255

Glu Glu Asp Glu Ala Ile Glu Val Glu Ala Thr Val
260 265

<210> 116
<211> 38
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (38)
<223> Xaa equals stop translation

<400> 116
Met Gly Cys Phe Pro Leu Trp Leu Val Thr Leu Ala Val Gly Asp Ala
1 5 10 15

Leu Pro Pro Thr Ala Cys Glu Leu Trp Gly Val Pro Ala Pro Pro Leu
20 25 30

His Leu Ala Glu Glu Xaa
35

<210> 117
<211> 122
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (122)
<223> Xaa equals stop translation

<400> 117
Met Gly Leu Trp Leu Gly Met Leu Ala Cys Val Phe Leu Ala Thr Ala
1 5 10 15

Ala Phe Val Ala Tyr Thr Ala Arg Leu Asp Trp Lys Leu Ala Ala Glu
20 25 30

Glu Ala Lys Lys His Ser Gly Arg Gln Gln Gln Gln Arg Ala Glu Ser
35 40 45

Thr Ala Thr Arg Pro Gly Pro Glu Lys Ala Val Leu Ser Ser Val Ala
50 55 60

Thr Gly Ser Ser Pro Gly Ile Thr Leu Thr Thr Tyr Ser Arg Ser Glu

65 70 75 80

Cys His Val Asp Phe Phe Arg Thr Pro Glu Glu Ala His Ala Leu Ser
85 90 95

Ala Pro Thr Ser Arg Leu Ser Val Lys Gln Leu Val Ile Arg Arg Gly
100 105 110

Ala Ala Leu Gly Ala Ala Ser Ala His Xaa
115 120

<210> 118
<211> 34
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (34)
<223> Xaa equals stop translation

<400> 118
Met Ile Gln Thr Phe Pro Ala Tyr Leu Cys Leu Pro Leu Phe Tyr Val
1 5 10 15

Leu Asp Leu Ala Leu Ala Ser Ala Pro Val Leu Ser His Ser Ala Leu
20 25 30

Leu Xaa

<210> 119
<211> 178
<212> PRT
<213> Homo sapiens

<400> 119
Met Gln Asn Asp Phe Gly Gln Val Trp Arg Trp Val Lys Glu Asp Ser
1 5 10 15

Ser Tyr Ala Asn Val Gln Asp Gly Phe Asn Gly Asp Thr Pro Leu Ile
20 25 30

Cys Ala Cys Arg Arg Gly His Val Arg Ile Val Ser Phe Leu Leu Arg
35 40 45

Arg Asn Ala Asn Val Asn Leu Lys Asn Gln Lys Glu Arg Thr Cys Leu
50 55 60

His Tyr Ala Val Lys Lys Lys Phe Thr Phe Ile Asp Tyr Leu Leu Ile
65 70 75 80

Ile Leu Leu Met Pro Val Leu Leu Ile Gly Tyr Phe Leu Met Val Ser

85 90 95

Lys Thr Lys Gln Asn Glu Ala Leu Val Arg Met Leu Leu Asp Ala Gly
100 105 110

Val Glu Val Asn Ala Thr Asp Cys Tyr Gly Cys Thr Ala Leu His Tyr
115 120 125

Ala Cys Glu Met Lys Asn Gln Ser Leu Ile Pro Leu Leu Leu Glu Ala
130 135 140

Arg Ala Asp Pro Thr Ile Lys Asn Lys His Gly Glu Ser Ser Leu Asp
145 150 155 160

Ile Ala Arg Arg Leu Lys Phe Ser Gln Ile Glu Leu Met Leu Arg Lys
165 170 175

Ala Leu

<210> 120
<211> 46
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (46)
<223> Xaa equals stop translation

<400> 120
Met Ile Leu Gln Ser Leu Leu Phe Leu Gln Arg Leu Leu Met Ile Ser
1 5 10 15

Thr Lys Pro Ala Val Val Leu Leu Trp Pro Leu Leu Lys Lys Val Glu
20 25 30

Asn Thr Leu Met Gln His Val His Pro Asn Leu Pro Ala Xaa
35 40 45

<210> 121
<211> 67
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (12)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (67)

<223> Xaa equals stop translation

<400> 121
Met Asn Leu Ser Ile Ile Leu Pro Asn Ser Phe Xaa His Leu Cys Asn
1 5 10 15

Phe Ser Leu Phe Leu Leu Pro Leu Pro Val Pro Ser Gln Pro Leu Ile
20 25 30

Cys Ser Gly Asn Tyr Gln Ser Ser Phe Cys His Tyr Arg Leu Ile Cys
35 40 45

Ile Phe Lys Glu Ile Tyr Ile His Gly Thr Ile His His Leu Cys Phe
50 55 60

Val Val Xaa
65

<210> 122
<211> 337
<212> PRT
<213> Homo sapiens

<400> 122
Met Glu Ile Arg Glu Glu Lys Lys Glu Asp Lys Val Glu Lys Leu Gln
1 5 10 15

Phe Glu Glu Glu Asp Phe Pro Ser Leu Asn Pro Glu Ala Gly Lys Gln
20 25 30

His Gln Pro Cys Arg Pro Ile Gly Thr Pro Ser Gly Val Trp Glu Asn
35 40 45

Pro Pro Ser Ala Lys Gln Pro Ser Lys Met Leu Val Ile Lys Lys Val
50 55 60

Ser Lys Glu Asp Pro Ala Ala Ala Phe Ser Ala Ala Phe Thr Ser Pro
65 70 75 80

Gly Ser His His Ala Asn Gly Asn Lys Leu Ser Ser Val Val Pro Ser
85 90 95

Val Tyr Lys Asn Leu Val Pro Lys Pro Val Pro Pro Pro Ser Lys Pro
100 105 110

Asn Ala Trp Lys Ala Asn Arg Met Glu His Lys Ser Gly Ser Leu Ser
115 120 125

Ser Ser Arg Glu Ser Ala Phe Thr Ser Pro Ile Ser Val Thr Lys Pro
130 135 140

Val Val Leu Ala Ser Gly Ala Ala Leu Ser Ser Pro Lys Glu Ser Pro
145 150 155 160

Ser Ser Thr Thr Pro Pro Ile Glu Ile Ser Ser Ser Arg Leu Thr Lys
165 170 175

Leu Thr Arg Arg Thr Thr Asp Arg Lys Ser Glu Phe Leu Lys Thr Leu
180 185 190

Lys Asp Asp Arg Asn Gly Asp Phe Ser Glu Asn Arg Asp Cys Asp Lys
195 200 205

Leu Glu Asp Leu Glu Asp Asn Ser Thr Pro Glu Pro Lys Glu Asn Gly
210 215 220

Glu Glu Gly Cys His Gln Asn Gly Leu Ala Leu Pro Val Val Glu Glu
225 230 235 240

Gly Glu Val Leu Ser His Ser Leu Glu Ala Glu His Arg Leu Leu Lys
245 250 255

Ala Met Gly Trp Gln Glu Tyr Pro Glu Asn Asp Glu Asn Cys Leu Pro
260 265 270

Leu Thr Glu Asp Glu Leu Lys Glu Phe His Met Lys Thr Glu Gln Leu
275 280 285

Arg Arg Asn Gly Phe Gly Lys Asn Gly Phe Leu Gln Ser Arg Ser Ser
290 295 300

Ser Leu Phe Ser Pro Trp Arg Ser Thr Cys Lys Ala Glu Phe Glu Asp
305 310 315 320

Ser Asp Thr Glu Thr Ser Ser Ser Glu Thr Ser Asp Asp Asp Ala Trp
325 330 335

Lys

<210> 123
<211> 69
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (69)
<223> Xaa equals stop translation

<400> 123
Met Lys Glu Ala Leu His Trp Ala Leu Phe Ser Met Gln Ala Thr Gly
1 5 10 15

His Val Leu Leu His Leu Leu Leu Pro Ala Ala Ala Pro Arg Cys His
20 25 30

Arg Gly Arg Ala Ser Pro Gln Gly Gln Gly Leu Ile Pro His Pro Asp

35 40 45

Leu Ser Glu Asp Thr Ala Val Lys Ala Gln Ala Leu Ala Phe Pro Ser
50 55 60

Glu Gly Leu Asp Xaa
65

<210> 124
<211> 77
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (60)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (77)
<223> Xaa equals stop translation

<400> 124
Met Asn Gly Gln Arg Met Asp Glu Leu Phe Val Leu Ile Arg Asp Gly
1 5 10 15

Phe Leu Leu Pro Thr Gly Leu Ser Ser Leu Ala Gln Leu Leu Leu Leu
20 25 30

Glu Ile Ile Glu Phe Arg Ala Ala Gly Trp Lys Thr Thr Pro Ala Ala
35 40 45

His Lys Tyr Tyr Tyr Ser Glu Ser Pro Thr Arg Xaa Pro Asp Gln Gly
50 55 60

Phe Leu Thr Ser Thr Gly Leu Ser Ser Thr His Leu Xaa
65 70 75

<210> 125
<211> 22
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (22)
<223> Xaa equals stop translation

<400> 125
Met Leu Leu Phe Leu Ile Leu Phe Phe Tyr Glu Lys Asn Gln Cys Gln
1 5 10 15

Ser Ala Asp Pro Leu Xaa
20

<210> 126
<211> 37
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (37)
<223> Xaa equals stop translation

<400> 126
Met Gly Lys Leu Leu Phe Pro Leu Leu Leu Ala Pro Phe Ser Pro Ile
1 5 10 15

Asn Lys Tyr Ile Leu His Phe Ala Arg Asp Gly Val Glu Glu Val Leu
20 25 30

Lys Phe Val Ser Xaa
35

<210> 127
<211> 62
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (62)
<223> Xaa equals stop translation

<400> 127
Met Leu Val Val Ala Val Ile Phe Leu His Gly Ala Gly Ala Met Asn
1 5 10 15

Tyr Leu Ile Ala Lys Ile Leu Glu Val Gln Gly Leu Arg Glu Val Pro
20 25 30

Cys Thr Tyr Asn Thr Arg Gly Ile Ala Pro Pro Gly Gly Asn Val Gly
35 40 45

Phe Glu Ala Ala Ser Val Val Asp Arg Pro Cys Gly Gln Xaa
50 55 60

<210> 128
<211> 46
<212> PRT
<213> Homo sapiens

<220>

<221> SITE
<222> (41)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (46)
<223> Xaa equals stop translation

<400> 128
Met Gly Phe Phe Glu Thr Ile Lys Leu Leu Leu Trp Val Val Leu Ile
1   5   10   15

Asp Cys Val Gly Val Gly Leu Leu Ile Ala Thr Leu Met Trp Phe Ile
            20   25   30

Ser Asn Lys Tyr Leu Val Lys Arg Xaa Glu Gln Arg Leu Xaa
        35   40   45

<210> 129
<211> 56
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (56)
<223> Xaa equals stop translation

<400> 129
Met Cys Ala Leu His Trp Leu His Trp Leu Ala Ser Trp Leu Cys Ser
1   5   10   15

Gln Pro Cys Leu Leu Leu Pro Ser Ser Pro Val Leu Cys Gln Ala Phe
            20   25   30

Ser Pro Ser Pro Val Ser Ser Pro Leu Arg Gln Ala Ile Ala Pro Ile
        35   40   45

Trp Leu Gly Arg His Arg Gln Xaa
    50   55

<210> 130
<211> 63
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (63)
<223> Xaa equals stop translation

<400> 130

Met Arg Glu Asp Pro Thr Trp Gly Arg Ser Leu Lys Ser Ser Leu Lys
1 5 10 15

Ile Leu Ser Asp Leu Ser Tyr Ser Leu Val Leu Trp Leu Thr Ala Ile
20 25 30

Leu Gly Leu Thr Ala Gln Lys Ser Gln Glu Lys Ser Gly Arg Ala Arg
35 40 45

Ile Gln Ser Ile Cys Ser Tyr Asn Val Ala Thr Ser Phe Ala Xaa
50 55 60

<210> 131
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (35)
<223> Xaa equals stop translation

<400> 131
Met Leu Ser Leu Met Ser His Leu His Val Gln Gln His Leu Ser Ser
1 5 10 15

Ile Leu Leu Ile Leu Ile Val Phe Ala Phe Leu Ser Asn Pro Phe Leu
20 25 30

Asn Gln Xaa
35

<210> 132
<211> 33
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (33)
<223> Xaa equals stop translation

<400> 132
Met Thr Arg Trp Leu Val Gln His His Thr Ser Leu Val Gln Val Leu
1 5 10 15

Ala Val Ser Phe Pro Ala Glu Gly Pro Gly Thr Glu Phe Pro Thr Ser
20 25 30

Xaa

<210> 133
<211> 118
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (118)
<223> Xaa equals stop translation

<400> 133
Met Gly Val Leu Cys Arg Ser Leu Ala Gly Leu Gly Gly Leu Ser Leu
1   5   10   15

Leu Gly Val Phe Cys Gly Gly Tyr Leu Met Ala Leu Ala Val Leu Ser
            20   25   30

Pro Cys Pro Pro Leu Val Gly Thr Ser Ala Gly Val Val Leu Val Val
        35   40   45

Leu Ser Trp Val Leu Cys Leu Gly Val Phe Ser Tyr Val Lys Val Ala
    50   55   60

Ala Ser Ser Leu Leu His Gly Gly Gly Arg Pro Ala Leu Leu Ala Ala
65   70   75   80

Gly Val Ala Ile Gln Val Gly Ser Leu Leu Gly Ala Val Ala Met Phe
                85   90   95

Pro Pro Thr Ser Ile Tyr His Val Phe His Ser Arg Lys Asp Cys Ala
            100   105   110

Asp Pro Cys Asp Ser Xaa
        115

<210> 134
<211> 146
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (146)
<223> Xaa equals stop translation

<400> 134
Met Leu Thr Arg Leu Val Leu Ser Ala His Leu Ser Ser Thr Thr Ser
1   5   10   15

Pro Pro Trp Thr His Ala Ala Ile Ser Trp Glu Leu Asp Asn Val Leu
            20   25   30

Met Pro Ser Pro Arg Ile Trp Pro Gln Val Thr Pro Thr Gly Arg Ser
        35   40   45

Ala Ser Val Arg Ser Glu Gly Asn Thr Ser Ser Leu Trp Asn Phe Ser
50 55 60

Ala Gly Gln Asp Val His Ala Ile Val Thr Arg Thr Cys Glu Ser Val
65 70 75 80

Leu Ser Ser Ala Val Tyr Thr His Gly Cys Gly Cys Val Arg Ser Ala
85 90 95

Thr Asn Ile Thr Cys Gln Ser Ser Gly Gln Gln Arg Gln Ala Ala Arg
100 105 110

Gln Glu Glu Glu Asn Ser Ile Cys Lys Ala His Asp Ser Arg Glu Gly
115 120 125

Arg Leu Gly Tyr Pro Leu Ser Ala His Gln Pro Gly Ser Gly Gly Pro
130 135 140

Asn Xaa
145

<210> 135
<211> 45
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (45)
<223> Xaa equals stop translation

<400> 135
Met Asn Arg Ile Leu Ser Tyr Leu Glu Thr Gly Phe Phe Ser Leu Pro
1 5 10 15

Leu Tyr Phe Phe Leu Thr Tyr Glu Leu His Val Pro Leu Met Lys Thr
20 25 30

Met Asn Trp Thr Cys Thr Thr Val His Val Ile Asp Xaa
35 40 45

<210> 136
<211> 134
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (114)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE
<222> (134)
<223> Xaa equals stop translation

<400> 136
Met Ala Leu Met Glu Val Asn Leu Leu Ser Gly Phe Met Val Pro Ser
1 5 10 15

Glu Ala Ile Ser Leu Ser Glu Thr Val Lys Lys Val Glu Tyr Asp His
20 25 30

Gly Lys Leu Asn Leu Tyr Leu Asp Ser Val Asn Glu Thr Gln Phe Cys
35 40 45

Val Asn Ile Pro Ala Val Arg Asn Phe Lys Val Ser Asn Thr Gln Asp
50 55 60

Ala Ser Val Ser Ile Val Asp Tyr Tyr Glu Pro Arg Arg Gln Ala Val
65 70 75 80

Arg Ser Tyr Asn Ser Glu Val Lys Leu Ser Ser Cys Asp Leu Cys Ser
85 90 95

Asp Val Gln Gly Cys Arg Pro Cys Glu Asp Gly Ala Ser Gly Ser His
100 105 110

His Xaa Ser Ser Val Ile Phe Ile Phe Cys Phe Lys Leu Leu Tyr Phe
115 120 125

Met Glu Leu Trp Leu Xaa
130

<210> 137
<211> 26
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (26)
<223> Xaa equals stop translation

<400> 137
Met Gln Lys Arg Glu Arg Lys Leu Tyr Val Ile Phe Leu Tyr Leu Ala
1 5 10 15

Phe Ile Leu Leu His Trp Gln Ser Gly Xaa
20 25

<210> 138
<211> 19
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (19)
<223> Xaa equals stop translation

<400> 138
Ile Pro Asn Glu Met Ala Gly Ser Ile Trp Pro Leu Gly Tyr Leu Ala
1 5 10 15

Thr Leu Xaa

<210> 139
<211> 93
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (93)
<223> Xaa equals stop translation

<400> 139
Met Leu Arg Cys Ala Trp Ala Leu Ala Pro Pro Val Pro Pro Pro Leu
1 5 10 15

Val Thr Asp Leu Pro Phe Phe Phe Thr Leu Ser Pro Phe Leu Phe Ala
20 25 30

Leu Glu Pro Pro Leu Pro Asp Leu Thr Asp Ser Ala Ser Met Ser Val
35 40 45

Ile Val Asp Arg Arg Ser Arg Gly Ser Asp Thr Asn Cys Trp Leu Leu
50 55 60

Asn Arg Arg Ser Lys His Pro Gly Ala Pro Arg Met Cys Thr Cys Lys
65 70 75 80

Ala Asn Ser Asn Lys Tyr Thr Ser Ser Leu Thr Asp Xaa
85 90

<210> 140
<211> 40
<212> PRT
<213> Homo sapiens

<400> 140
Met Arg Ala Asn Phe Arg Cys Trp Leu His Cys Thr Leu Tyr Leu Leu
1 5 10 15

Cys Ser Pro Pro Ser Asn Gln Gly Ser Cys Gln Cys Thr Pro His Val
20 25 30

Pro Trp Arg Ser Trp Cys Cys Glu
35 40

<210> 141
<211> 82
<212> PRT
<213> Homo sapiens

<400> 141
Met Ser Ala His Cys Asn Leu His Leu Pro Gly Ser Ser Asn Ser Pro
1 5 10 15

Thr Ser Ala Ser Gln Val Ala Gly Ile Thr Arg Glu Glu Ala Glu Gly
20 25 30

Gln Gly Gly Lys Gly Ile Gly Ser Gln Val His Gly Pro Leu Val Lys
35 40 45

Pro Pro Leu Leu Trp Gly Leu Arg Lys His Arg Gly Gly Val Ser Cys
50 55 60

Ser Ala Cys Pro His Ser Pro Ala Asn Asn Val Val Thr Ser Val Pro
65 70 75 80

Asn Leu

<210> 142
<211> 76
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (76)
<223> Xaa equals stop translation

<400> 142
Met Asn Met Cys Trp Gln Ile Pro Asn Phe Ile Leu Ile Gln Val Ser
1 5 10 15

Ser Glu Tyr Val His Ile Leu Ile Val Ile Val Thr Lys Thr Pro Gly
20 25 30

Val Gln Ser Gly Ser Cys Cys Ser Leu His Arg Lys Pro Met Pro Glu
35 40 45

Thr Thr Ser Val Ala Lys Glu Glu Gly Leu Ile Gly Cys Cys Ser Arg
50 55 60

Gly Asp Gly Ser Ser Val Ser Asn Pro Ser Leu Xaa
65 70 75

<210> 143
<211> 92
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (86)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 143
Met Arg Met Pro Ser His Thr His Ala Arg Phe Val Leu Phe Tyr Leu
1 5 10 15

Ile Leu Arg Asn Arg Ser Gly Gly Val Leu Pro Gly Cys Ser Asp Pro
20 25 30

Glu Gly Ser Gln Glu Ser Pro Gly Leu Gln Lys Ser Pro Pro Thr Gly
35 40 45

Ser Glu Ala Ser Leu Ser Trp Cys Ile Gln Thr Ala His Ser Arg Leu
50 55 60

Trp Ala Leu Thr Leu Gln Ile Pro Glu Ser Pro Pro Gly Leu Pro Ala
65 70 75 80

Leu Gly Pro Val Pro Xaa Ser Ser Lys Gly Gly Arg
85 90

<210> 144
<211> 23
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (18)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (23)
<223> Xaa equals stop translation

<400> 144
Met Leu Pro Lys Pro Gln Leu Ser Val Leu Thr Leu Thr Val Ala Leu
1 5 10 15

Ser Xaa Ile Pro Gly Thr Xaa
20

<210> 145
<211> 40
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (40)
<223> Xaa equals stop translation

<400> 145
Met Glu Met Met Met Val Val Met Gly Cys Val Gln Gly Pro Gly Glu
1   5   10   15

Gly Cys Ser Gly Lys Met Gly Lys Lys Pro Arg Pro Trp Pro Leu Val
            20   25   30

Ser Tyr Ser Ile Thr His Leu Xaa
        35   40

<210> 146
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (35)
<223> Xaa equals stop translation

<400> 146
Met Leu Leu Tyr Gln Ile Asn Ile Pro Phe Ser Phe Ala Leu Ser Val
1   5   10   15

Leu Leu Ser Leu Cys Trp Pro His Gln His Tyr Tyr Pro Cys Tyr Ile
            20   25   30

Ser Phe Xaa
        35

<210> 147
<211> 34
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (34)
<223> Xaa equals stop translation

<400> 147
Met Cys Val Cys Val Phe Ser Phe Cys Leu Phe Cys Leu Phe Val Phe
1   5   10   15

Gly Met Val Leu Thr Val Leu Leu Cys His Pro Gly Trp Ser Ala Val
20 25 30

Val Xaa

<210> 148
<211> 51
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (51)
<223> Xaa equals stop translation

<400> 148
Met Leu Ile Phe Cys Gly Glu Tyr Trp Tyr Phe Cys Phe Asn Leu Leu
1 5 10 15

Trp Val Val Val Pro Tyr Lys Phe Ser Phe Leu Ser Phe Gly Ser Val
20 25 30

Ile Gln Ile Cys Pro Thr Ser Val Pro Pro Ile Gly Gln Ser Gly Ile
35 40 45

Trp Val Xaa
50

<210> 149
<211> 83
<212> PRT
<213> Homo sapiens

<400> 149
Met Arg Phe Leu Lys Leu Phe Ser His Asn Ile Leu Ile Gln Leu Lys
1 5 10 15

Ile Ile Leu Lys Leu Lys Val Ser Ser Val Leu Pro Ser Val Lys Ser
20 25 30

Leu Lys Asp Glu Arg Ile Ile Phe Ile Phe Gln Val Ser Leu Asn Lys
35 40 45

Val Leu Ser Pro Cys Leu Arg Phe Tyr Pro Gln Arg Thr Ala Thr Phe
50 55 60

Leu Ser Cys Gln Ile Glu Phe Val Gln Gln Leu Arg Asn Thr Gly Lys
65 70 75 80

Ile Gln Asn

<210> 150
<211> 47
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (47)
<223> Xaa equals stop translation

<400> 150
Met Lys Glu Lys Gln Val Tyr His Ile Ser Lys Ile Lys Glu Glu Tyr
1 5 10 15

Ser Ile Leu Ile Cys Leu Leu Ile Val Lys Met Ser Phe Pro Gln Ile
20 25 30

Ala Pro Ile Gln Phe Lys Arg Lys His Ser Thr Lys Ile Gln Xaa
35 40 45

<210> 151
<211> 49
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (49)
<223> Xaa equals stop translation

<400> 151
Met Trp Asp Gln Arg Pro Thr Lys Gly Thr Gln Asp Phe Gln Leu Leu
1 5 10 15

Leu Leu Pro Gly Ile Cys Ser Ser Phe Ala Leu Leu Leu Asn Ala Leu
20 25 30

Pro Phe Pro Ala Pro Ser Pro Ser Ile Gly Thr Cys Leu Cys Ala Ser
35 40 45

Xaa

<210> 152
<211> 77
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (73)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (77)
<223> Xaa equals stop translation

<400> 152
Met Gln Trp Val His Ile Ala Glu Thr Gly Asn Glu Lys Phe Ser Phe
1   5   10   15

Phe Leu Phe Phe Phe Cys Gly Gly Trp Gly Gln Ser Leu Thr Leu Ser
            20   25   30

Pro Arg Gln Glu Cys Ser Gly Ala Ile Ser Ala His Cys Asn Leu Pro
        35   40   45

Pro Pro His Leu Gln Val Gln Ala Ile Leu Val Pro Pro Pro Pro Glu
    50   55   60

Gln Leu Ala Leu Gln Val His Ala Xaa Thr Leu Gly Xaa
65   70   75

<210> 153
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (35)
<223> Xaa equals stop translation

<400> 153
Met Phe Tyr Asp Val Gln Gly Pro Ser His Ser Ser Glu Met Cys Phe
1   5   10   15

Phe Val Phe Phe Phe Val Cys Leu Phe Leu Phe Leu Met Asn Glu Ser
            20   25   30

Lys Gly Xaa
        35

<210> 154
<211> 65
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (65)
<223> Xaa equals stop translation

<400> 154
Met Val Leu Leu Leu Trp Arg Leu Phe Phe Pro Val Gly Leu Met Arg
1   5   10   15

Ile Ala Gln Pro Leu Gly His Leu Ile Lys His Arg Glu Thr Tyr Ser
            20   25   30

Leu Arg His Trp Cys Leu His Thr Gln Val Met Leu Gly His Gly Asp
        35   40   45

Glu Thr Ala Pro Leu Leu Ile Phe Leu Lys Lys Pro Ser Cys His Ile
    50   55   60

Xaa
65

<210> 155
<211> 85
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (85)
<223> Xaa equals stop translation

<400> 155
Met Ser Ile Gln Val Leu Cys Pro Leu Phe Cys Phe Ala Ser Phe Phe
1   5   10   15

Ile Leu Gly Ser Arg Gly Glu Cys Ala Gly Phe Tyr Thr His Val Leu
            20   25   30

Gln Asp Pro Arg Ala Trp Ala Ser Asn Asp Pro Ala Thr Gln Val Val
        35   40   45

Asn Ile Val Pro Asn Arg Glu Phe Ser Thr Leu Ala Leu Leu Leu Pro
    50   55   60

Pro His Phe Trp Asn Pro Trp Cys Pro Leu Phe Pro Cys Cys Ala Met
65   70   75   80

Cys Pro Gln Cys Xaa
                85

<210> 156
<211> 3
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (3)

<223> Xaa equals stop translation

<400> 156
Met Ser Xaa

<210> 157
<211> 42
<212> PRT
<213> Homo sapiens

<400> 157
Met Ala Gly Arg Gly Arg Gly Arg Val Ala Ser Ser Trp Val Gly Gly
1 5 10 15

Thr Gly Pro Thr Cys Cys Gly Cys Lys Trp Pro Gly Gln Leu Thr Glu
20 25 30

His Leu Leu Phe Ala Asp Pro Thr Leu Arg
35 40

<210> 158
<211> 32
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (32)
<223> Xaa equals stop translation

<400> 158
Met Ser Arg Ala Asn Lys Glu Ile Met Leu Leu Leu Pro Ala Asp Val
1 5 10 15

Pro Leu Val Tyr Ser Val Val Ser Val Gly Arg Val Thr Leu Arg Xaa
20 25 30

<210> 159
<211> 47
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (47)
<223> Xaa equals stop translation

<400> 159

Met Trp Asn Phe Ser Cys Ser Thr Ser Ile Cys Glu Tyr Gly Phe Leu
1   5   10   15

Lys Phe Leu Val Leu Tyr Leu Leu Ser Thr Ser Met Ser Ser Pro Leu
            20   25   30

Ile Gly Pro Glu Pro His Ser Pro Thr Lys Cys Lys Ile Lys Xaa
        35   40   45

<210> 160
<211> 159
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (159)
<223> Xaa equals stop translation

<400> 160
Met Val Phe Val Val Leu Leu Pro Glu Met Ile Pro Leu Thr Ala Glu
1   5   10   15

Glu Gly Gly Gly Trp Lys Lys Ser Arg Ser Asp Pro Lys Thr Leu Pro
            20   25   30

Val Gln Ala Phe Val Phe Lys Cys Gln Ala Trp Gly Pro Arg Arg Arg
        35   40   45

Arg Glu Gly Leu Pro Trp Asp Ser Ser Lys Leu Ser Pro Leu Ser Ser
    50   55   60

Thr Arg Leu Thr Thr Cys Ser Pro Pro Pro Thr Ser Gly Arg Gly Leu
65   70   75   80

Gln Gly Thr Gln Glu Ala Ala Pro Trp Thr Pro Gly Pro Ser Pro Thr
                85   90   95

Lys Pro Ser Val Pro Lys Ala Pro Asp Pro Glu Leu Ala Arg Thr Met
            100   105   110

Gln Ala Gly Leu Leu Trp Val Leu Ala Glu Pro Ala Thr Asn Gly Gly
        115   120   125

Arg Glu Gly Arg Arg Ser Leu Thr Phe Ser Gln Asn Lys Pro Arg Arg
    130   135   140

Asn Pro Arg Lys Ala Glu Val Leu Phe Phe Ala Asn Pro Val Xaa
145   150   155

<210> 161
<211> 90
<212> PRT

```
<213> Homo sapiens

<220>
<221> SITE
<222> (90)
<223> Xaa equals stop translation

<400> 161
Met Val Val Pro Ala Asp Ser Gly Gly Leu Pro Arg Arg Thr Glu Lys
  1               5                  10                  15

Leu Leu Cys Val Met Leu Leu Leu Leu Glu Arg Met Ala Leu Cys Pro
             20                  25                  30

Val Leu Asp Val His Thr His Leu Gly Cys Ile Ile Cys Val Ala Cys
         35                  40                  45

Gln Pro Val Arg Thr Val Leu Ser Leu Leu Thr Ala Ser Ile Gln Glu
     50                  55                  60

Gly Ser Arg Leu Ser Gly His Phe Gln Thr Leu Pro His Gln Thr Asp
 65                  70                  75                  80

Thr Thr Phe His Lys Gly Ser Lys Leu Xaa
                 85                  90


<210> 162
<211> 64
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (13)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 162
Met Thr Leu Ile Thr Pro Ala Arg Ile Thr Leu Thr Xaa Gly Asn Lys
  1               5                  10                  15

Ser Trp Ser Ser Thr Ala Val Ala Ala Ala Leu Glu Leu Val Asp Pro
             20                  25                  30

Pro Gly Cys Arg Asn Ser Ala Arg Asp Arg Cys Met His Thr Pro Leu
         35                  40                  45

Cys Val Cys Met Cys Val Cys Val Cys Val Cys Arg Gly Ile Leu Val
     50                  55                  60


<210> 163
```

<211> 146
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (146)
<223> Xaa equals stop translation

<400> 163
Met Ser Leu Phe Cys Leu Lys Leu Leu Ser Gly Cys Leu Trp Leu Ser
1 5 10 15

Gly Ser Glu Pro His His Gly Leu Gly Phe Leu Leu Trp Pro Leu Ala
20 25 30

Phe Ala Ser Cys Ser Ile Leu Ile Leu Asn Tyr Ala Lys Pro Phe Leu
35 40 45

Asn Pro Ala Pro Cys Ser Leu Cys Leu Glu Leu Pro Ser Gln Ala Phe
50 55 60

Leu Cys Arg Ser Phe Ser Ser His Leu Leu Ser Glu Pro Ser Leu Val
65 70 75 80

Thr Pro Phe His His Pro Val Cys Phe Leu Pro Ile Ile Trp Phe Pro
85 90 95

Trp Arg Leu Met Ser Val Ser Pro Gln Trp Asn Val Gly Leu Met Ala
100 105 110

Gln Ala His Arg Gly His Cys Cys Val Gln Gly Ser Val Arg Met Pro
115 120 125

Arg Cys Ala Trp Met Trp Arg Trp Pro Ala Gly Trp Gly Cys His Leu
130 135 140

Ala Xaa
145

<210> 164
<211> 69
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (69)
<223> Xaa equals stop translation

<400> 164
Met Gly Thr Glu Gln Ser Leu Gly Tyr Arg Val Gln Gly Leu Leu Leu
1 5 10 15

Val Leu Ser Leu His Val Ser Gln Arg Gly Leu Cys Gly Ser Leu Pro
20 25 30

Pro Ser Met Ser Ser Glu Glu Arg Lys Gln Arg Pro Trp Ser Ser Gln
35 40 45

Tyr Gly Glu His Cys Val Pro Asp Thr Pro Leu Arg Val Lys Val Arg
50 55 60

Arg His Ile Leu Xaa
65

<210> 165
<211> 89
<212> PRT
<213> Homo sapiens

<400> 165
Met Arg Glu Thr Thr Pro Met Ile Gln Leu Pro Pro Ser Gly Ser Pro
1 5 10 15

Phe Ile Cys Gly Asp Tyr Glu Tyr Tyr His Leu Arg Glu Ile Leu Asn
20 25 30

Gly Ser Thr Asp Pro Asn His Ser Thr Ala Leu Arg Tyr Leu Ile Ile
35 40 45

Lys Leu Pro Lys Val Lys Gly Lys Glu Arg Ile Leu Lys Ile Ala Arg
50 55 60

Glu Lys Lys Gln Ile Thr Cys Asn Gly Ala Pro Ile Cys Leu Ala Ala
65 70 75 80

Asp Val Ser Val Glu Thr Leu Leu Val
85

<210> 166
<211> 88
<212> PRT
<213> Homo sapiens

<400> 166
Met His Phe Trp Thr Gly Pro Arg Phe Gln Leu Gly Leu Ala Gly Val
1 5 10 15

Pro Ala Ala Gln Phe Glu Thr Ser His Ile Glu Ser Arg Ala Arg Ser
20 25 30

Arg Ala Cys Gly Lys Phe Leu Gly Phe Cys Ser Ser Arg Thr Val Pro
35 40 45

Ser Ala Trp Cys Glu Ala Leu Met Glu Pro Ala Val Ile Gly Tyr Glu
50 55 60

Thr Lys Ser Leu Pro Ile His Gly Cys Pro Phe Ile His Trp His Arg
65 70 75 80

Thr Pro Gly Thr Asn Glu Gly Asp
85

<210> 167
<211> 37
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (37)
<223> Xaa equals stop translation

<400> 167
Met Leu Asp Pro Ala Ala Ser Gly Thr Phe Arg Ala Leu Leu Leu Leu
1 5 10 15

Ser His Pro Phe Leu Asp Trp Ser Leu Ser Asp Pro His Cys Glu Ser
20 25 30

Leu Asn Gln Lys Xaa
35

<210> 168
<211> 34
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (34)
<223> Xaa equals stop translation

<400> 168
Met Ser His Asn Ile Gln Pro Leu Phe Ser Phe Leu Thr Leu Leu Ser
1 5 10 15

Tyr Phe Leu Phe His Phe Leu Ser Leu Pro Ser Ser Phe Phe Pro Asn
20 25 30

Tyr Xaa

<210> 169
<211> 36
<212> PRT
<213> Homo sapiens

<400> 169
Met Pro Ser Leu Pro Ile Arg Val Thr Lys Phe Ser Glu Ile Gly Asn
1 5 10 15

Trp Gln Leu Lys Ala Val Ser Thr Thr Arg Phe Leu Leu Pro Leu Lys
20 25 30

Lys Asn His Phe
35

<210> 170
<211> 57
<212> PRT
<213> Homo sapiens

<400> 170
Met Leu Leu Lys Ser Thr Gly Ser Phe Leu Glu Phe Gly Leu Gln Glu
1 5 10 15

Ser Cys Ala Glu Phe Trp Thr Ser Ala Asp Asp Ser Ser Ala Ser Asp
20 25 30

Glu Ile Arg Leu Glu Leu Cys Phe Leu Ser Pro Ser Thr Ser Tyr Leu
35 40 45

Val Val Ser Phe Leu Met Val Arg Ser
50 55

<210> 171
<211> 45
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (45)
<223> Xaa equals stop translation

<400> 171
Met Tyr Val Lys Ala Ser Ala Val Thr Val Ser Arg Asp Glu Ala Leu
1 5 10 15

Thr Pro Cys Leu Pro Asp Pro His Trp Asn Ala Pro Phe Ala Arg His
20 25 30

Leu Leu Gln Pro Ser Cys Ser Phe Leu Glu Phe Pro Xaa
35 40 45

<210> 172
<211> 96
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (96)
<223> Xaa equals stop translation

<400> 172
Met Leu Ser Glu Thr Pro His Ala Arg Arg Gly Arg Ala Phe Leu Thr
1 5 10 15

Asp Ser Leu Pro Met Val Ile Pro Ser Leu Leu Leu Pro Pro Pro Gly
20 25 30

Arg Ala Ser Leu Ala Glu Pro Thr Leu Arg Ser Val Lys Gly Gln Pro
35 40 45

Leu Thr Leu Ser Gln His Met Glu Asp Leu Ala Val Ser Arg Glu Asn
50 55 60

Cys Ser His Tyr Arg Val Gln Leu Cys Pro Pro Ala Pro Ala Pro Ser
65 70 75 80

Ala Pro Arg Leu Thr Leu Met Ala Leu Ser Cys Ser Ser Leu Pro Xaa
85 90 95

<210> 173
<211> 49
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (42)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (49)
<223> Xaa equals stop translation

<400> 173
Met Trp Asp Thr Phe Val Arg Asp Arg Asp Phe Ser Ala Tyr Leu Phe
1 5 10 15

Leu His Leu Leu Pro Pro Leu Ser Ala Cys Gly Leu Asn Ala Ser Leu
20 25 30

Tyr Thr Ala Thr Pro Ile Val Trp Val Xaa His Thr Ser Pro Gln Asp
35 40 45

Xaa

<210> 174
<211> 50
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (50)
<223> Xaa equals stop translation

<400> 174
Met Val Arg Ser Ser Ser His Phe Lys Phe Phe Leu Met Leu Phe Thr
1 5 10 15

Ser Thr Leu Gln Asp Val Gly His Thr Ser His Pro Ser Ala Gln Pro
20 25 30

Ser Ser Arg Leu Ser Asp Ser Pro Leu Ile Cys Leu Ile Asn Arg Gln
35 40 45

Val Xaa
50

<210> 175
<211> 61
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (61)
<223> Xaa equals stop translation

<400> 175
Met Thr Pro Gly Val Gly Ala Glu Pro Arg Gly Glu Gly Cys Lys Gly
1 5 10 15

Lys Ala Val Arg Gly Leu Gly Gly Glu Arg Val Ser Pro Val Leu Leu
20 25 30

Val Leu His Leu Arg Ser Pro Ser Pro Val Glu Gly Glu Gln Ser Gln
35 40 45

Arg Gln Trp Gly Val Gln Phe Trp Asn Leu Glu Glu Xaa
50 55 60

<210> 176
<211> 40
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (15)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (36)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (40)
<223> Xaa equals stop translation

<400> 176
Ile Leu Gly Phe Ser Phe Ala Val Gly Glu Gly Lys Trp Gly Xaa Phe
1 5 10 15

Cys Leu Leu Val Pro Gly Ile Met Leu His Ile Ile His Leu Leu Ser
20 25 30

His Leu Ile Xaa Pro Asn Pro Xaa
35 40

<210> 177
<211> 53
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (53)
<223> Xaa equals stop translation

<400> 177
Met Pro Leu Asp Leu Leu Phe Leu Ile Thr Tyr Phe Leu Leu Ser Val
1 5 10 15

Ile Leu Lys Val Leu Tyr Ile Asp Ala Pro Gly His Leu Gly Met Pro
20 25 30

Ile Ser Leu Cys Ser Ser Ala Val Val Trp Val Lys Val Asp Leu Val
35 40 45

Ser Glu Lys Gly Xaa
50

<210> 178
<211> 41
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (41)
<223> Xaa equals stop translation

<400> 178
Met Ser Val Leu Ser Gly Phe Leu Phe Ile Val Val Val Cys Cys Tyr
1 5 10 15

Cys Cys Phe Val Ala Arg Leu Gln Leu Thr Lys Tyr Glu Phe Lys Asn
20 25 30

Cys Val Val Ile Phe Arg Asp Leu Xaa
35 40

<210> 179
<211> 105
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (105)
<223> Xaa equals stop translation

<400> 179
Met Glu Arg Asp Thr Arg Glu Lys Cys Leu Trp Ser Leu Pro Tyr Pro
1 5 10 15

Lys Leu Leu Cys Asn Leu Leu Ala Ser His Phe Leu Ser Ile Leu Ser
20 25 30

Phe Phe Ile Tyr Ser Ile Gly Phe Leu Asp Leu Val Val Ser Asn Thr
35 40 45

Leu Pro Val Phe Gln Phe Asp Val Thr Phe Tyr Pro Val Thr Lys Phe
50 55 60

Ile Phe Gln Lys His Ser Met Leu Cys His Thr Ala Asn Leu Val Asn
65 70 75 80

Val Pro Asp Met Val Trp Leu Cys Pro His Pro Asn Leu Ile Leu Asn
85 90 95

Cys Ser Ser His Asn Pro His Met Xaa
100 105

<210> 180
<211> 40
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (40)
<223> Xaa equals stop translation

<400> 180
Met Asp Tyr Glu Val Ile Ser Gln Asn Val Arg Lys Arg Tyr Arg Ala
1 5 10 15

Leu Glu Leu Leu Tyr Leu Leu Leu Asn Leu Asn Ile Thr Ala Thr Asn
20 25 30

Lys Gly Tyr Gln Asp Lys Val Xaa
35 40

<210> 181
<211> 25
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (25)
<223> Xaa equals stop translation

<400> 181
Met Ile Tyr Phe Leu Leu Leu Leu Pro Glu Ala Gln Gly Glu Phe Ser
1 5 10 15

Ser Ile Phe Thr Val Arg Thr Trp Xaa
20 25

<210> 182
<211> 54
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (13)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (54)
<223> Xaa equals stop translation

<400> 182
Met Cys Pro Pro Ser Gln Arg Ala Pro Thr His Leu Xaa Cys Pro Trp
1 5 10 15

Val Asp Pro Gly Pro Val Val Leu Gly Leu Ser Leu Trp Val Leu Ala
20 25 30

Gly Gly Met Gly Glu Gly Gly Glu Gln Leu Pro Ala Pro Leu Leu Cys
35 40 45

Gly Ser Ser Phe Phe Xaa
50

<210> 183
<211> 66
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (50)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (66)
<223> Xaa equals stop translation

<400> 183
Met Leu Leu Asn Thr Ser Phe Thr Arg Glu Ile Ile Ile Ser Gln Arg
1 5 10 15

Glu Ser Asn Trp Leu Val Leu Leu Leu Leu Leu Phe Phe Pro Val Ile
20 25 30

Cys Phe Ile Glu Arg Ser Leu Cys Gly Gly Thr Asp Phe Leu Asn Thr
35 40 45

Leu Xaa His Thr His Thr Tyr Thr Pro Ser Ile Tyr Gly Ala Met His
50 55 60

Arg Xaa
65

<210> 184
<211> 27
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (27)
<223> Xaa equals stop translation

<400> 184
Met Ile His Leu Ser Arg Phe Tyr Leu Leu Leu Ile Met Leu Pro His
1 5 10 15

Val Leu Phe Phe Thr Gly Asp Leu His Ser Xaa

20 25

<210> 185
<211> 24
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (24)
<223> Xaa equals stop translation

<400> 185
Met Phe Pro Phe Pro Phe Phe His Leu Val Ile Leu Gly Phe Leu Leu
1 5 10 15

Leu His Ser Phe Leu Pro Pro Xaa
20

<210> 186
<211> 42
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (42)
<223> Xaa equals stop translation

<400> 186
Met Ser Gln Thr Leu Val Ala Leu Pro Glu Arg Asn Glu Asn Ala Gln
1 5 10 15

Pro His Pro Cys Thr Leu Cys Ser Phe Leu Phe Asn Thr Glu Glu Pro
20 25 30

Glu Trp Arg Gly Pro Ala Gly Leu Gln Xaa
35 40

<210> 187
<211> 223
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (75)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (146)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (159)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 187
Met Val Pro Arg Thr Ser His Thr Ala Ala Phe Leu Ser Asp Thr Lys
1 5 10 15

Asp Arg Gly Pro Pro Val Gln Ser Gln Ile Trp Arg Ser Gly Glu Lys
20 25 30

Val Pro Phe Val Gln Thr Tyr Ser Leu Arg Ala Phe Glu Lys Pro Pro
35 40 45

Gln Val Gln Thr Gln Ala Leu Arg Asp Phe Glu Lys His Leu Asn Asp
50 55 60

Leu Lys Lys Glu Asn Phe Ser Leu Lys Leu Xaa Ile Tyr Phe Leu Glu
65 70 75 80

Glu Arg Met Gln Gln Lys Tyr Glu Ala Ser Arg Glu Asp Ile Tyr Lys
85 90 95

Arg Asn Thr Glu Leu Lys Val Glu Val Glu Ser Leu Lys Arg Glu Leu
100 105 110

Gln Asp Lys Lys Gln His Leu Asp Lys Thr Trp Ala Asp Val Glu Asn
115 120 125

Leu Asn Ser Gln Asn Glu Ala Glu Leu Arg Arg Gln Phe Glu Glu Arg
130 135 140

His Xaa Glu Thr Glu His Val Tyr Glu Leu Leu Glu Asn Lys Xaa Gln
145 150 155 160

Leu Leu Gln Glu Glu Ser Arg Leu Ala Lys Asn Glu Ala Ala Arg Met
165 170 175

Ala Ala Leu Val Glu Ala Glu Lys Glu Cys Asn Leu Glu Leu Ser Glu
180 185 190

Lys Leu Lys Gly Val Thr Lys Asn Trp Glu Asp Val Pro Gly Asp Gln
195 200 205

Val Lys Pro Asp Gln Tyr Thr Glu Ala Leu Ala Gln Arg Asp Lys
210 215 220

<210> 188
<211> 239
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (91)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (162)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (175)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 188
```
Met Glu Gln Thr Trp Thr Arg Asp Tyr Phe Ala Glu Asp Asp Gly Glu
  1               5                  10                  15

Met Val Pro Arg Thr Ser His Thr Ala Ala Phe Leu Ser Asp Thr Lys
             20                  25                  30

Asp Arg Gly Pro Pro Val Gln Ser Gln Ile Trp Arg Ser Gly Glu Lys
         35                  40                  45

Val Pro Phe Val Gln Thr Tyr Ser Leu Arg Ala Phe Glu Lys Pro Pro
     50                  55                  60

Gln Val Gln Thr Gln Ala Leu Arg Asp Phe Glu Lys His Leu Asn Asp
 65                  70                  75                  80

Leu Lys Lys Glu Asn Phe Ser Leu Lys Leu Xaa Ile Tyr Phe Leu Glu
                 85                  90                  95

Glu Arg Met Gln Gln Lys Tyr Glu Ala Ser Arg Glu Asp Ile Tyr Lys
            100                 105                 110

Arg Asn Thr Glu Leu Lys Val Glu Val Glu Ser Leu Lys Arg Glu Leu
        115                 120                 125

Gln Asp Lys Lys Gln His Leu Asp Lys Thr Trp Ala Asp Val Glu Asn
    130                 135                 140

Leu Asn Ser Gln Asn Glu Ala Glu Leu Arg Arg Gln Phe Glu Glu Arg
145                 150                 155                 160

His Xaa Glu Thr Glu His Val Tyr Glu Leu Leu Glu Asn Lys Xaa Gln
                165                 170                 175

Leu Leu Gln Glu Glu Ser Arg Leu Ala Lys Asn Glu Ala Ala Arg Met
            180                 185                 190

Ala Ala Leu Val Glu Ala Glu Lys Glu Cys Asn Leu Glu Leu Ser Glu
        195                 200                 205
```

Lys Leu Lys Gly Val Thr Lys Asn Trp Glu Asp Val Pro Gly Asp Gln
210 215 220

Val Lys Pro Asp Gln Tyr Thr Glu Ala Leu Ala Gln Arg Asp Lys
225 230 235

<210> 189
<211> 228
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (66)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (127)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (131)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (141)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 189
Ile Arg His Glu Leu Leu Pro Ala Leu His Leu Gln Ala His Asp Ala
1 5 10 15

Ala Tyr Asn Leu Leu Phe Phe Ala Ser Gly Gly Gly Lys Phe Asn Tyr
20 25 30

Gln Gly Thr Lys Arg Trp Leu Glu Asp Asn Leu Asp His Thr Gly Glu
35 40 45

Arg Pro Arg Val Gly Val Gly Val Pro Arg Trp Trp Cys Arg Gly Glu
50 55 60

Ala Xaa Arg Pro Arg Gly Cys His Gly Gly Ser Gln Glu Ala Gln Arg
65 70 75 80

Glu Gly Arg Gly Pro Leu Pro Gly Pro His Pro Pro Arg Gln Leu Ser
85 90 95

Val Ser Cys Arg Leu Gln Pro Ala Ser Gly Gln Cys Gly Leu Arg Ala
100 105 110

Val Pro Gly His Arg Gly Pro Gly Gln Gln Pro Ala Pro Ala Xaa Val
115 120 125

Arg Pro Xaa Arg Glu Gly Thr Leu Gln His Ala Phe Xaa Arg Glu Leu
130 135 140

Glu Thr Val Ala Ala His Gln Phe Pro Glu Val Arg Phe Ser Met Val
145 150 155 160

His Lys Arg Ile Asn Leu Ala Glu Asp Val Leu Ala Trp Glu His Glu
165 170 175

Arg Phe Ala Ile Arg Arg Leu Pro Ala Phe Thr Leu Ser His Leu Glu
180 185 190

Ser His Arg Asp Gly Gln Arg Ser Ser Ile Met Asp Val Arg Ser Arg
195 200 205

Val Asp Ser Lys Thr Leu Ile Arg Leu Pro Gln Pro Pro Lys Val Leu
210 215 220

Gly Leu Arg Val
225

<210> 190
<211> 40
<212> PRT
<213> Homo sapiens

<400> 190
Ile Tyr Leu Asn Ile Gln Val Val Arg Gly Gln Arg Lys Val Ile Cys
1 5 10 15

Leu Leu Lys Glu Gln Ile Ser Asn Glu Gly Glu Asp Lys Ile Phe Leu
20 25 30

Ile Asn Lys Leu His Ser Ile Tyr
35 40

<210> 191
<211> 27
<212> PRT
<213> Homo sapiens

<400> 191
Glu Arg Lys Glu Arg Glu Glu Arg Ser Arg Val Gly Thr Thr Glu Glu
1 5 10 15

Ala Ala Ala Pro Pro Ala Leu Leu Thr Asp Glu
20 25

<210> 192

<211> 7
<212> PRT
<213> Homo sapiens

<400> 192
Arg His Glu Met Glu Asn Thr
1 5

<210> 193
<211> 53
<212> PRT
<213> Homo sapiens

<400> 193
Arg Lys Leu Ser Thr Gly Pro Phe Ser Ala Cys Lys Pro Arg Ala Thr
1 5 10 15

Cys Cys Phe Thr Ser Cys Tyr Leu Gln Gln Leu Leu Asp Ala Thr Glu
20 25 30

Asp Gly His Pro Pro Lys Gly Lys Ala Ser Ser Leu Ile Pro Thr Cys
35 40 45

Leu Lys Ile Leu Gln
50

<210> 194
<211> 29
<212> PRT
<213> Homo sapiens

<400> 194
Thr Ser Cys Tyr Leu Gln Gln Leu Leu Asp Ala Thr Glu Asp Gly His
1 5 10 15

Pro Pro Lys Gly Lys Ala Ser Ser Leu Ile Pro Thr Cys
20 25

<210> 195
<211> 25
<212> PRT
<213> Homo sapiens

<400> 195
Cys Cys Gly Ala Lys Arg Ile Met Lys Glu Ala Leu His Trp Ala Leu
1 5 10 15

Phe Ser Met Gln Ala Thr Gly His Val
20 25

<210> 196

<211> 196
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (13)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (15)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (91)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (126)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 196
Pro Pro Ala Gly Ala Thr Ser Pro Gly Arg Ile Ile Xaa Pro Xaa Ser
1 5 10 15

Ala Val Leu Ile Pro Ser Pro Val Lys Ser Tyr Arg Gly Trp Leu Val
20 25 30

Met Gly Glu Pro Ser Arg Glu Glu Tyr Lys Ile Gln Ser Phe Asp Ala
35 40 45

Glu Thr Gln Gln Leu Leu Lys Thr Ala Leu Lys Asp Pro Gly Ala Val
50 55 60

Asp Leu Glu Lys Val Ala Asn Val Ile Val Asp His Ser Leu Gln Asp
65 70 75 80

Cys Val Phe Ser Lys Glu Ala Gly Arg Met Xaa Tyr Ala Ile Ile Gln
85 90 95

Ala Glu Ser Lys Gln Ala Gly Gln Ser Val Phe Arg Arg Gly Leu Leu
100 105 110

Asn Arg Leu Gln Gln Glu Tyr Gln Ala Arg Glu Gln Leu Xaa Ala Arg
115 120 125

Ser Leu Gln Gly Trp Val Cys Tyr Val Thr Phe Ile Cys Asn Ile Phe
130 135 140

Asp Tyr Leu Arg Val Asn Asn Met Pro Met Met Ala Leu Val Asn Pro
145 150 155 160

Val Tyr Asp Cys Leu Phe Arg Leu Ala Gln Pro Asp Ser Leu Ser Lys
165 170 175

Glu Glu Glu Val Asp Cys Leu Val Leu Gln Leu His Arg Val Gly Glu
180 185 190

Gln Leu Glu Lys
195

<210> 197
<211> 24
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (6)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (8)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 197
Pro Gly Arg Ile Ile Xaa Pro Xaa Ser Ala Val Leu Ile Pro Ser Pro
1 5 10 15

Val Lys Ser Tyr Arg Gly Trp Leu
20

<210> 198
<211> 25
<212> PRT
<213> Homo sapiens

<400> 198
Lys Gln Ala Gly Gln Ser Val Phe Arg Arg Gly Leu Leu Asn Arg Leu
1 5 10 15

Gln Gln Glu Tyr Gln Ala Arg Glu Gln
20 25

<210> 199
<211> 21
<212> PRT
<213> Homo sapiens

<400> 199
Tyr Asp Cys Leu Phe Arg Leu Ala Gln Pro Asp Ser Leu Ser Lys Glu
1 5 10 15

Glu Glu Val Asp Cys
20

<210> 200
<211> 127
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (19)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 200
Met Lys Arg Thr Ser Val Asn Pro Gln Thr Leu Cys Glu Ala Arg Pro
1 5 10 15

Ala Gly Xaa Ser Gln Gln Pro Leu Ser Leu Asp Ser Glu Ala Pro Arg
20 25 30

Gly Gly Val Ala Pro Pro Arg Leu Gln Gly Pro Pro Pro His Gln Arg
35 40 45

Val His Leu Thr Leu Glu Cys Thr Thr His Pro Thr Val Gly Lys Ala
50 55 60

Ser Val Leu Gly Pro Cys Leu Leu Leu Leu Ser Cys Pro Arg Ala Pro
65 70 75 80

Ala Gly Pro Pro Pro Pro Pro His Ser Arg Val Arg Ala Gly Gly Cys
85 90 95

Arg Pro Trp Ala Arg Arg Glu Gly His Cys Arg Pro Leu Gly Ala Asp
100 105 110

Thr Asp Thr Ser Arg Ile Cys His Gly Arg Arg Pro Phe Ser Leu
115 120 125

<210> 201
<211> 76
<212> PRT
<213> Homo sapiens

<400> 201
Met Ser Leu Pro Ala Ala Pro Ala Gly Arg Leu Ser Pro Leu Tyr Trp
1 5 10 15

Arg Ser Ser Asn Thr Arg Ser Gln Leu Ser Leu Leu Trp Glu Leu Gly
20 25 30

His Phe Phe Thr Arg Cys Cys Arg Arg Pro His Pro Asn Pro His Leu
35 40 45

Pro Ala Leu Ser Val Cys Arg Cys His Ile Leu His Lys Ile Met Leu
50 55 60

Trp Glu Pro Ser Ser Pro Leu Leu Pro Ala Leu Pro
65 70 75

<210> 202
<211> 86
<212> PRT
<213> Homo sapiens

<400> 202
Met Thr Ser Pro Gly Gln Gly Arg Ala Gly Arg Arg Gly Asp Glu Gly
1 5 10 15

Ser His Asn Met Ile Leu Cys Lys Ile Trp Gln Arg His Thr Leu Arg
20 25 30

Ala Gly Arg Trp Gly Leu Gly Trp Gly Arg Arg Gln His Arg Val Lys
35 40 45

Lys Cys Pro Ser Ser His Ser Lys Glu Ser Cys Asp Arg Val Phe Glu
50 55 60

Leu Leu Gln Tyr Lys Gly Glu Ser Arg Pro Ala Gly Ala Ala Gly Arg
65 70 75 80

Asp Ile Ile Trp Phe Pro
85

<210> 203
<211> 17
<212> PRT
<213> Homo sapiens

<400> 203
Pro Ser Leu Arg Gly Pro Lys Ala Gly Ala Pro Pro Arg Trp Arg Pro
1 5 10 15

Leu

<210> 204
<211> 25
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (7)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 204
Asn Leu Val Asp Pro Pro Xaa Cys Arg Asn Ser Ala Arg Glu Thr Leu
1 5 10 15

Lys Leu Gly Arg Val Glu Val Ser Ile
20 25

<210> 205
<211> 7
<212> PRT
<213> Homo sapiens

<400> 205
Lys Ala Gly Ala Pro Pro Arg
1 5

<210> 206
<211> 6
<212> PRT
<213> Homo sapiens

<400> 206
Cys Arg Asn Ser Ala Arg
1 5

<210> 207
<211> 109
<212> PRT
<213> Homo sapiens

<400> 207
Gln Asp Ser Arg Lys Met Leu Pro Ser Thr Ser Val Asn Ser Leu Val
1 5 10 15

Gln Gly Asn Gly Val Leu Asn Ser Arg Asp Ala Ala Arg His Thr Ala
20 25 30

Gly Ala Lys Arg Tyr Lys Tyr Leu Arg Arg Leu Phe Arg Phe Arg Gln
35 40 45

Met Asp Phe Glu Phe Ala Ala Trp Gln Met Leu Tyr Leu Phe Thr Ser
50 55 60

Pro Gln Arg Val Tyr Arg Asn Phe His Tyr Arg Lys Gln Thr Lys Asp
65 70 75 80

Gln Trp Ala Arg Asp Asp Pro Ala Phe Leu Val Leu Leu Ser Ile Trp
85 90 95

Leu Cys Val Ser Thr Ile Gly Phe Gly Phe Val Leu Asp
100 105

<210> 208
<211> 117
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (2)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 208
Asn Xaa Gln Ser Arg Asp Tyr Asp Val Glu Trp Gly Tyr Ala Phe Asp
1 5 10 15

Val His Leu Asn Ala Phe Tyr Pro Leu Leu Val Ile Leu His Phe Ile
20 25 30

Gln Leu Phe Phe Ile Asn His Val Ile Leu Thr Asp Thr Phe Ile Gly
35 40 45

Tyr Leu Val Gly Asn Thr Leu Trp Leu Val Ala Val Gly Tyr Tyr Ile
50 55 60

Tyr Val Thr Phe Leu Gly Tyr Ser Ala Leu Pro Phe Leu Lys Asn Thr
65 70 75 80

Val Ile Leu Leu Tyr Pro Phe Ala Pro Leu Ile Leu Leu Tyr Gly Leu
85 90 95

Ser Leu Ala Leu Gly Trp Asn Phe Thr His Thr Leu Cys Ser Phe Tyr
100 105 110

Lys Tyr Arg Val Lys
115

<210> 209
<211> 45
<212> PRT
<213> Homo sapiens

<400> 209
Ser Val Asn Ser Leu Val Gln Gly Asn Gly Val Leu Asn Ser Arg Asp
1 5 10 15

Ala Ala Arg His Thr Ala Gly Ala Lys Arg Tyr Lys Tyr Leu Arg Arg
20 25 30

Leu Phe Arg Phe Arg Gln Met Asp Phe Glu Phe Ala Ala
35 40 45

<210> 210
<211> 23

<212> PRT
<213> Homo sapiens

<400> 210
Val Ile Leu Thr Asp Thr Phe Ile Gly Tyr Leu Val Gly Asn Thr Leu
1 5 10 15

Trp Leu Val Ala Val Gly Tyr
20

<210> 211
<211> 16
<212> PRT
<213> Homo sapiens

<400> 211
Gly Trp Asn Phe Thr His Thr Leu Cys Ser Phe Tyr Lys Tyr Arg Val
1 5 10 15

<210> 212
<211> 20
<212> PRT
<213> Homo sapiens

<400> 212
Pro Met Val Leu Lys Leu Lys Asp Trp Pro Pro Gly Glu Asp Phe Arg
1 5 10 15

Asp Met Met Pro
20

<210> 213
<211> 16
<212> PRT
<213> Homo sapiens

<400> 213
Tyr Phe Val Arg Pro Asp Leu Gly Pro Lys Met Tyr Asn Ala Tyr Gly
1 5 10 15

<210> 214
<211> 9
<212> PRT
<213> Homo sapiens

<400> 214
Asn Ser Ala Arg Glu Asp Gly Gln Pro
1 5

<210> 215
<211> 8
<212> PRT
<213> Homo sapiens

<400> 215
Leu Asn Leu Ala Ser Arg Leu Pro
1 5

<210> 216
<211> 153
<212> PRT
<213> Homo sapiens

<400> 216
Val Lys Pro Asp Pro Pro Arg Ala Pro Gly Glu Asn Glu Asp Ser Ser
1 5 10 15

Val Pro Glu Thr Pro Asp Asn Glu Arg Lys Ala Ser Ile Ser Tyr Phe
20 25 30

Lys Asn Gln Arg Gly Ile Gln Tyr Ile Asp Leu Ser Ser Asp Ser Glu
35 40 45

Asp Val Val Ser Pro Asn Cys Ser Asn Thr Val Gln Glu Lys Thr Phe
50 55 60

Asn Lys Asp Thr Val Ile Ile Val Ser Glu Pro Ser Glu Asp Glu Glu
65 70 75 80

Ser Gln Gly Leu Pro Thr Met Ala Arg Arg Asn Asp Asp Ile Ser Glu
85 90 95

Leu Glu Asp Leu Ser Glu Leu Glu Asp Leu Lys Asp Ala Lys Leu Gln
100 105 110

Thr Leu Lys Glu Leu Phe Pro Gln Arg Ser Asp Asn Asp Leu Leu Lys
115 120 125

Val Ile Phe Ile Gly Tyr Cys Ser Cys Asn Asp Asp Lys Ile Ser Pro
130 135 140

Ala Phe Ser Ala Ile Val Ser Ser Gly
145 150

<210> 217
<211> 17
<212> PRT

<213> Homo sapiens

<400> 217
Lys Asp Ala Lys Leu Gln Thr Leu Lys Glu Leu Phe Pro Gln Arg Ser
1 5 10 15

Asp

<210> 218
<211> 16
<212> PRT
<213> Homo sapiens

<400> 218
Lys Asp Thr Val Ile Ile Val Ser Glu Pro Ser Glu Asp Glu Glu Ser
1 5 10 15

<210> 219
<211> 16
<212> PRT
<213> Homo sapiens

<400> 219
Glu Asp Ser Ser Val Pro Glu Thr Pro Asp Asn Glu Arg Lys Ala Ser
1 5 10 15

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later that at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

INTERNATIONAL SEARCH REPORT	International application No. PCT/US98/15949

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C07H 19/00; C12P 19/34; C12Q 1/68
US CL : 536/22.1; 435/91.41; 435/6
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/22.1; 435/91.41; 435/6

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

cas, APS, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	STEWART et al. Growth, Differentiation, and survival: Multiple Physiological Functions for Insulin-like Growth-factors. Physiological Reviews. October 1996, Vol. 76, No.4, pages 1005-1026, see entire article.	1-23
Y	JOHNSTONE et al. Immunochemistry in Practice. London: Blackwell Scientific Publications. second edition, 1987, page 30.	1-23

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26 OCTOBER 1998	16 DEC 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer EGGERTON CAMPBELL Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/15949

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SKONIER et al. Big-h3: A Transforming Growth Factor-B-Responsive Gene Encoding a Secreted Protein That Inhibits Cell Attachment In Vitro and Suppresses the Growth of CHO Cells in Nude Mice. DNA and Cell Biology. 1994, Vol. 13, No. 6, pages 571-584, see entire article.	1-23
Y	XUAN et al. Recombinant PSP94 Demonstrates Similar Linear Epitote Structure as Natural PSP94 Protein. J. Cellular Biochemistry. 1996, Vol. 63, pages 61-73, see entire article.	1-23
Y	EP 0 559 428 A2 (ONO PHARMACEUTICAL CO, LTD.l) 08 September 1993, see entire document.	1-23
Y	LEWIS et al Rescue, Expression, and Analysis of a Neutralizing Human Anti-Hepatitis A Virus Monoclonal Antibody. J. Immunology. 01 September 1993, Vol. 151, No. 5, pages 2829-2838, see entire article.	1-23